

09/914464

FORM PTO-1390 (REV. 12-2001)		U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER CL000895USNAT	
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371					
INTERNATIONAL APPLICATION NO. PCT/US00/28888		INTERNATIONAL FILING DATE 08 November 2000		PRIORITY DATE CLAIMED 24 November 1999	
TITLE OF INVENTION PRIMARY SEQUENCE OF THE SHRIMP WHITE SPOT BACILLIFORM VIRUS ...					
APPLICANT(S) FOR DO/EO/US PE CORPORATION (NY) et al.					
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:					
<p>1. <input checked="" type="checkbox"/> This is a <b>FIRST</b> submission of items concerning a filing under 35 U.S.C. 371.</p> <p>2. <input type="checkbox"/> This is a <b>SECOND</b> or <b>SUBSEQUENT</b> submission of items concerning a filing under 35 U.S.C. 371.</p> <p>3. <input type="checkbox"/> This is an express request to begin national examination procedures (35 U.S.C. 371(f)). The submission must include items (5), (6), (9) and (21) indicated below.</p> <p>4. <input type="checkbox"/> The US has been elected by the expiration of 19 months from the priority date (Article 31).</p> <p>5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2))</p> <p>a. <input type="checkbox"/> is attached hereto (required only if not communicated by the International Bureau).</p> <p>b. <input type="checkbox"/> has been communicated by the International Bureau.</p> <p>c. <input checked="" type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US).</p> <p>6. <input type="checkbox"/> An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).</p> <p>a. <input type="checkbox"/> is attached hereto.</p> <p>b. <input type="checkbox"/> has been previously submitted under 35 U.S.C. 154(d)(4).</p> <p>7. <input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))</p> <p>a. <input type="checkbox"/> are attached hereto (required only if not communicated by the International Bureau).</p> <p>b. <input type="checkbox"/> have been communicated by the International Bureau.</p> <p>c. <input checked="" type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired.</p> <p>d. <input type="checkbox"/> have not been made and will not be made.</p> <p>8. <input type="checkbox"/> An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371 (c)(3)).</p> <p>9. <input type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).</p> <p>10. <input type="checkbox"/> An English language translation of the annexes of the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).</p> <p>Items 11 to 20 below concern document(s) or information included:</p> <p>11. <input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.</p> <p>12. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.</p> <p>13. <input type="checkbox"/> A <b>FIRST</b> preliminary amendment.</p> <p>14. <input type="checkbox"/> A <b>SECOND</b> or <b>SUBSEQUENT</b> preliminary amendment.</p> <p>15. <input type="checkbox"/> A substitute specification.</p> <p>16. <input type="checkbox"/> A change of power of attorney and/or address letter. previously submitted</p> <p>17. <input type="checkbox"/> A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U.S.C. 1.821 - 1.825.</p> <p>18. <input type="checkbox"/> A second copy of the published international application under 35 U.S.C. 154(d)(4).</p> <p>19. <input type="checkbox"/> A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4).</p> <p>20. <input checked="" type="checkbox"/> Other items or information: copy of Grant of Petition dated 20 December 2001</p>					

U.S. APPLICATION NO (if known, see 37 CFR 1.5) <b>09/914,464</b>		INTERNATIONAL APPLICATION NO. <b>PCT/US00/28888</b>		ATTORNEY'S DOCKET NUMBER <b>CL000895USNAT</b>	
---	--	--	--	--	--

<b>21. <input type="checkbox"/> The following fees are submitted:</b> <b>BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)):</b> Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO..... <b>\$1040.00</b> International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO ..... <b>\$890.00</b> International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO ..... <b>\$740.00</b> International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4) ..... <b>\$710.00</b> International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4) ..... <b>\$100.00</b> <b>ENTER APPROPRIATE BASIC FEE AMOUNT =</b>				<b>CALCULATIONS PTO USE ONLY</b>  <div style="border: 1px solid black; height: 100px; width: 100%;"></div>	
Surcharge of <b>\$130.00</b> for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				previously \$ paid not yet \$ provided	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE	\$	
Total claims	51 - 20 =	31	x \$18.00	\$	558
Independent claims	10 - 3 =	7	x \$84.00	\$	588
MULTIPLE DEPENDENT CLAIM(S) (if applicable)				+	\$ 280
<b>TOTAL OF ABOVE CALCULATIONS =</b>				\$	1426
<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.				+	\$ -0-
<b>SUBTOTAL =</b>				\$	1426
Processing fee of <b>\$130.00</b> for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(f)).				\$	-0-
<b>TOTAL NATIONAL FEE =</b>				\$	1426
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property +				\$	-0-
<b>TOTAL FEES ENCLOSED =</b>				\$	1426
				Amount to be refunded:	\$
				charged:	\$

a. ☐ A check in the amount of \$ \_\_\_\_\_ to cover the above fees is enclosed.

b. ☒ Please charge my Deposit Account No. 50-0970 in the amount of \$ 1426.00 to cover the above fees. A duplicate copy of this sheet is enclosed.

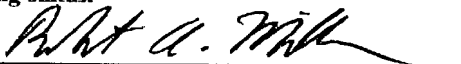
c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 50-0970. A duplicate copy of this sheet is enclosed.

d. ☐ Fees are to be charged to a credit card. **WARNING:** Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

**NOTE:** Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137 (a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:  
 Celera Genomics  
 Attn: Robert A. Millman  
 Chief Intellectual Property Counsel  
 Customer Number 25748  
 45 West Gude Drive, C2-4#20  
 Rockville, MD 20850-1159

  
 SIGNATURE  
 Robert A. Millman  
 NAME  
 36,217  
 REGISTRATION NUMBER



**PRIMARY NUCLEOTIDE SEQUENCE OF THE SHRIMP WHITE SPOT  
BACILLIFORM VIRUS (WSBV), DISCOVERY SYSTEMS CONTAINING THIS  
SEQUENCE AND DETECTION KITS AND ANTIVIRAL TARGETS FOR  
DETECTION AND CONTROLLING SHRIMP VIRUS OUTBREAK AND SPREAD**

**RELATED APPLICATION**

The present application claims priority to Chinese patent application No. 99124717.5, filed November 24, 1999.

**FIELD OF THE INVENTION**

The present invention is in the field of genomic discovery systems. The present invention specifically provides the complete shrimp white spot bacilliform virus (WSBV) genome and isolated fragments thereof in a form that is commercially useful, including detection kits, antiviral agents, reagents such as nucleic acid arrays, and computer-based systems.

**BACKGROUND OF THE INVENTION**

The shrimp and prawn (hereafter collectively referred to as shrimp) industry is a rapid growth worldwide industry worth billions of dollars. Worldwide, the shrimp industry relies on both the harvesting of wild shrimp and aquaculture, which is the controlled farming of fish, shellfish, and plants. Aquaculture, particularly aquaculture of shrimp, is growing rapidly due to increasing consumer demand for shrimp and other seafood. Aquaculture has been expanding at an annual rate far surpassing the growth of livestock meat, capture fisheries, and agricultural production. The aquaculture industry delivers high-quality protein for human and animal consumption and provides a substantial source of income and employment, particularly for developing countries. Aquaculture accounts for nearly 20 percent of the world's harvest of fish, shellfish, and seaweeds. The total worldwide value of giant tiger prawn production is the greatest of any aquaculture species. Aquaculture of giant tiger prawn significantly contributes to many Asian and Latin American economies, where the majority of giant tiger prawn production occurs. Shrimp accounted for approximately a quarter of the overall value of Asian fish exports in 1996. In the United States, harvesting and processing shrimp, including both aquaculture and harvesting wild shrimp, is a \$3 billion dollar a year industry that employs over 11,000 people. Furthermore, shrimp aquaculture in the U.S. has

the potential to become a high-growth business. The risk of viral diseases to cultured shrimp is the primary obstacle to the growth of the shrimp aquaculture industry.

Shrimp viral disease is a major worldwide concern of the shrimp industry. Both aquaculture and wild shrimp are vulnerable to viral infection, which can lead to devastating economic consequences. Furthermore, shrimp viruses may affect other crustaceans such as crabs and crayfish. Drastic declines in the populations of wild shrimp or other crustaceans due to viral disease can also dramatically affect other species in the food chain that depend upon shrimp for food and can lead to severe ecological consequences.

Major pathogenic shrimp viruses include White Spot Bacilliform Virus (WSBV), Infectious Hypodermal and Hematopoietic Virus (IHHNV), Taura Syndrome Virus (TSV), and Yellow Head Virus (YHV). IHHNV and TSV are endemic throughout South and Central America, while WSBV and YHV are endemic throughout Asia. All U.S. shrimp species are susceptible to infection and disease from one or more of these four viruses. Susceptibility of U.S. species of shrimp to these viruses may lead to restrictions on the importation of foreign shrimp into the U.S.

Past incidents of viral outbreaks illustrate the devastating affects that a viral outbreak can have on the shrimp industry. An outbreak of IHHNV in 1987 in the Gulf of California shrimp fishery reduced shrimp to levels that could not support commercial harvests until 1994. Outbreaks in 1995 and 1996 on U.S. shrimp farms caused a 50 to 95 percent loss of production at affected farms. Shrimp exports from China to the U.S. dropped 75% between 1990 and 1995 due to infection by WSBV.

WSBV is regarded as one of the most highly pathogenic viruses of penaeid shrimp. No uniform name exists for WSBV. It is also known as White Spot Syndrome Virus (WSSV), Prawn White Spot Bacilliform Virus (PWSBV), White Spot Baculiform Virus (WSBV), Baculoviral Hypodermal and Hematopoietic Necrosis Virus (HHNBV), Rod-shaped Nuclear Virus of *Penaeus japonicus* (RV-PJ), Systemic Ectodermal and Mesodermal Bacilliform Virus (SEMBV). Other acronyms include WSV, WSDV, and LNBV. The virus is a non-occluded, circular, double-stranded DNA bacilliform virus with a genome of approximately 300kb. WSBV virions are enveloped nucleocapsids with bacilliform morphology and a tail-like extension at one end.

White Spot Syndrome, caused by WSBV, is also known by such names as Red Disease, China Virus Disease, and Shrimp Explosive Epidemic Disease. Infected shrimp display rapid reduction in food consumption and lethargy. Gross observations include a loose cuticle and a red color to the entire body and appendages along with small subcutaneous

white spots. Histological examination reveals prominent intranuclear inclusion bodies in the cuticular epithelium, subcutis, and connective tissues. Cumulative mortality rates reach 100% within 3 to 10 days of the onset of clinical signs. No significant resistance to WSBV has been reported. All native U.S. species of shrimp are susceptible to WSBV infection under experimental conditions. WSBV is widely spread throughout most of the shrimp growing regions of Asia and the Indo-Pacific, including China, Japan, Korea, Thailand, Indonesia, Taiwan, Vietnam, Malaysia, and India. Lethal outbreaks of WSBV virus have recently been recorded in Texas and South Carolina. Furthermore, the virus has been shown to infect other crustaceans including amphipods, ostracods, swimming crabs, crayfish, copepods, and shore flies. The possibility exists that these organisms could act as a reservoir through which further shrimp infection, or infection of other species, can occur.

In view of the serious economic and ecological risks posed to the worldwide shrimp industry and shrimp populations by viruses, particularly WSBV, a strong need exists for antiviral agents and detection systems. Detection systems should be highly specific, rapid, and sensitive. To facilitate development of antiviral agents and detection systems, knowledge of the complete genomic sequence and protein encoding sequences of WSBV is needed. Prior to the present invention, very few reports on WSBV genomic sequences existed, and only a small fraction of the entire WSBV genome had been sequenced. To date, only six WSBV sequenced have been patented, published or stored in public genome databases, such as Genbank. All sequences to date are short sequences ranging in length from 420 bp to 2424 bp. J.S. Kim and others from Korea have sequenced 2424bp (wsu 92007, 1997) and 420bp (wsu 89843, 1997); K.Mitsuo and others from Japan have sequenced two fragments, 1447bp (PN JP 1997201196-A/2) and 1461bp (PN JP 1997201196-A/1) in length; Chufang Luo et.al. from Taiwan sequenced 1461bp (PMU50923, 1996); L.M. Nunan et. al. from the United States has reported 868bp sequence in *J. Virological Methods* (1997(63): p193-201). These known sequences are no more than 10kb in length all together. In addition, these sequences are randomly sequenced with no systematic analysis; therefore determining sequence function is difficult. Since the complete genome of WSBV is more than 300kb in length, the analysis of the complete genomic DNA sequence and it's complete structure, the determination of the expressed sequences, and prediction of the functions of encoded proteins are all new scientific achievements. These achievements are the basis for the present invention. The present invention is directed to providing the complete primary nucleotide sequence of WSBV and isolated fragments thereof, protein encoding sequences of WSBV,

and antiviral agents and detection systems based on the nucleotide and protein encoding sequences provided by the present invention.

#### DNA Viruses

Generally, transcription of a DNA virus genome occurs in the nucleus of the host cell, utilizing host cell polymerases and other host enzymes for viral mRNA synthesis and viral replication. Viral gene transcription is modulated by the interaction of specific DNA-binding proteins with promoter and enhancer elements in the viral genome. Commonly, the viral promoter and enhancer elements are similar in sequence to those of the host cell in order to allow the host cell's transcriptional activation factors and DNA-dependent RNA polymerase to bind the viral control elements. Cells from different tissues or species express different DNA-binding proteins, and this is a major factor in determining which species, and which cells and tissues of that species, that the virus can infect.

Viruses, in general, depend on the host cell ribosomes, transfer RNA (tRNA), and mechanisms of posttranslational modification to produce their proteins. Generally, viral mRNA encoding non-structural viral proteins, such as DNA-binding proteins and enzymes, are transcribed first. These are followed by late viral gene products encoding structural proteins.

Viruses utilize various methods to promote preferential translation of their viral mRNA over host cell mRNA. In some instances, concentration of viral mRNA in the host cell is so large that it occupies most of the cell's ribosomes, thereby preventing translation of host cell mRNA. Viruses may inhibit synthesis and/or induce degradation of the cell's nucleic acids. Many viruses increase the permeability of the host cell membrane, thereby reducing the ribosomal affinity for most cellular mRNA.

Viral DNA replication begins at a unique sequence in the genome called the origin of replication, or *ori*. The *ori* is recognized by viral or host nuclear factors and DNA-dependent DNA polymerase. Viral DNA synthesis is semi-conservative and a primer is required by the DNA polymerase to initiate synthesis of the new DNA molecule.

#### Viral Screening Tests and Antiviral Agents

Viral screening tests and detection kits, such as nucleic acid arrays, can be developed based on either nucleic acids or polypeptides provided by the present invention. A nucleic acid probe to a virus specific nucleotide sequence, or an antibody to a virus specific protein,

is introduced into contact with a sample, such as a sample of shrimp cells, whereby the presence of the virus is detected using an assay system.

Antiviral agents, either nucleic acid or protein-based, directly interfere with viral function or preferably, interfere with viral replication to stop or prevent spread of the virus in a population, such as in a population of shrimp. Knowledge of the nucleic acid and protein sequences of the virus allows antiviral agents to be designed to attack a number of viral targets necessary for viral replication or function, such as viral encoded enzymes or structural proteins. Attachment of the virus to the host cell is the first step in viral replication and is mediated by the interaction of a viral attachment protein and a host cell surface receptor. This interaction can be blocked by neutralizing antibodies, which bind to and coat the virion, or receptor antagonists which are peptide or carbohydrate analogues of the viral attachment protein and competitively block the interaction of the virus with the cell. Agents can be designed that bind to the viral attachment protein and prevent penetration of the virus into the cytoplasm or nucleus of the host cell and/or uncoating of the virus. These agents thereby prevent the virus from delivering its genome into the host cell. Viral mRNA expression and utilization can be targeted with anti-viral agents. Antisense oligonucleotides can be designed to bind to newly transcribed viral RNA and thereby prevent the viral RNA from being processed to mRNA in the nucleus, delivered to the cytoplasm, and bound to the ribosome. Many antiviral drugs are nucleoside analogues, which inhibit viral polymerases. Viral polymerases are often less specific for substrate than are host polymerases, therefore the viral polymerase will often bind a nucleotide analogue with a modified base and/or sugar several hundredfold better than the host enzyme. Antiviral drugs can therefore be preferentially incorporated into the viral genome. DNA viruses, such as WSBV, are particularly susceptible to these types of drugs due to the extent and rapid rate of nucleotide incorporation during viral replication. Inhibition of posttranslational modification of viral proteins, such as phosphorylation, may also inhibit viral replication.

#### **SUMMARY OF THE INVENTION**

The present invention is based on the sequencing and assembly of the WSBV genome. The present invention provides the primary nucleotide sequence of the WSBV genome (SEQ ID NO: 1) and predicted transcript sequences (SEQ ID NOS: 2, 4, 6...280, 282, 284, 286-293: See the Sequence Listing and the Figure Sheets for both the genomic and transcript sequences) and polymorphic sites on these transcripts summarized in Table 1

hereinafter, and protein encoded sequence produced from each of the genes found in the WSBV genome. This information is provided in the form of sequences and annotation information and can be used to generate computer based discovery systems, nucleic acid detection reagents and kits such as nucleic acid arrays, protein based detection kits, and antiviral targets.

The present invention provides these nucleotide sequences of the WSBV genome, and representative fragments thereof, in a form that can be used, analyzed, and commercialized. For example, the present invention provides the nucleic acid sequences as contiguous strings of primary sequences in a form readable by computers, such as recorded on computer readable media, e.g., magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. The present invention specifically provides a Sequence Listing in computer readable form stored on such media. Such compositions are useful in the discovery of drug and antiviral targets.

The present invention further provides systems, particularly computer-based systems that contain the primary sequence information of the present invention stored in data storage means. Such systems are designed to identify commercially important fragments of the WSBV genome.

Another embodiment of the present invention is directed to isolated fragments, and collections of fragments, of the WSBV genome. The fragments of the WSBV genome include, but are not limited to, fragments that encode peptides, hereinafter open reading frames (ORFs) and fragments that modulate the expression of an operably linked ORF, hereinafter expression modulating fragments (EMFs). The ORFs are provided in the Sequence Listing and in Figure 3.

The present invention further includes kits, such as nucleic acid arrays, detection reagents and microfluidic devices, that comprise one or more fragments of the WSBV genome of the present invention, particularly ORFs. The kits, such as arrays, can be used to track the expression of many genes, even all genes, or rationally selected subsets thereof, contained in the WSBV genome.

The identification of the entire coding set of sequences from the genome of WSBV will be of great value to all laboratories working with this organism and for a variety of commercial and ecological purposes. Many fragments of the WSBV genome will be immediately identified by similarity searches against protein and nucleic acid databases and

by identifying structural motifs present in protein domains and will be of immediate value to WSBV researchers and for commercial value for controlling WSBV infection in shrimp populations. A specific example concerns viral envelope proteins, many of which interact with host cells. Proteins of this family can readily be configured into screens and assays for detecting chemical modulators of the protein activity. The biological significance of this and other families of proteins for controlling viral replication is well known. Many of the known antiviral agents modulate the activity of these types of proteins. The WSBV genome will allow one to identify all potential antiviral targets.

The present invention is further directed to isolated WSBV proteins encoded by the ORFs of the present invention. A variety of methodologies known in the art can be utilized to obtain any one of the proteins of the present invention. The amino acid sequence can be synthesized using commercially available peptide synthesizers. In an alternative method, the viral protein can be purified from cells infected with the virus.

The invention further provides antibodies that selectively bind one of the WSBV proteins encoded by the present invention. Antibodies have use in viral detection and control and can be generated using the protein encoding sequences provided by the present invention. Such antibodies include both monoclonal and polyclonal antibodies, and fragments thereof. The invention further provides hybridomas capable of producing the above-described monoclonal antibodies.

The present invention provides methods of identifying WSBV in a test sample, such as a sample of shrimp. Such methods comprise incubating cells extracted from the test sample with one or more of the antibodies or probes based on the nucleic acid sequences provided by the present invention under conditions that allow a skilled artisan to determine if the test sample contains the ORF or product produced therefrom.

Using the isolated proteins of the present invention, the present invention further provides methods of obtaining and identifying agents capable of binding to a protein encoded by one of the ORFs of the present invention. Specifically, such agents include antibodies, peptides, carbohydrates, pharmaceutical agents and the like. Such methods comprise the steps of contacting an agent with an isolated protein encoded by one of the ORFs of the present invention and determining whether the agent binds to said protein.

#### **DESCRIPTION OF THE FIGURE SHEETS**

Figure 1 provides a block diagram of a computer system 102 that can be used to implement the computer-based systems of the present invention.

Figure 2 (Sheets 1-40) provides the primary genomic sequence of WSBV.

Figure 3 (Sheets 1-160) provides:

- 1) the predicted transcript sequence of the WSBV gene and starting ATG site (SEQ ID NOS: 2, 4, 6, 8 . . . 280, 282, 284, 286-293);
- 2) the predicted protein sequence of the WSBV gene (SEQ ID NOS: 3, 5, 7, 9 . . . 281, 283, 285);
- 3) results of a BLAST query run using default parameters that shows proteins producing significant alignments with the predicted WSBV protein sequence of the present invention.
- 4) comments
- 5) TaqMan primer/probe sets. Oligonucleotide sequences useful as primers and/or probes for amplifying and/or screening for the WSBV genes provided by the present invention.

## **DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS**

### General Description

The present invention is based on the sequencing and assembly of the WSBV genome. In this process, the primary nucleotide sequence of 5795 nucleic acid fragments was determined. These fragments were assembled into a single contiguous sequence of 305,107 bp. After assembly, the sequences were analyzed with various computer packages and compared with all external data sources. The result of this analysis was the identification of 150 predicted genes/transcripts contained in the WSBV genome. The present invention provides the genomic nucleic acid sequence of WSBV (SEQ ID NO: 1), see Figure 2, Sheets 1-40, as well as the predicted gene structure of all 150 identified genes (SEQ ID NOS: 2, 4, 6...280, 282, 284, 286-293 ) and polymorphic sites on these transcripts summarized in Table 1, and predicted amino acid sequences of all of the encoded proteins (SEQ ID NOS: 3, 5, 7...281, 283, 285), see Figure 3, sheets 1-160.



The nucleotide sequences of the present invention, or representative fragments thereof, are provided in a form that can be readily used, analyzed, and interpreted by a skilled

artisan. In one embodiment, the sequences are provided as contiguous strings of primary sequence information corresponding to the nucleotide sequences provided in the figures.

As used herein, a "representative fragment of the nucleotide sequence provided herein" refers to any portion of these sequences that are not presently represented within a publicly available database. Preferred representative fragments of the present invention are WSBV open reading frames and expression modulating fragments (ORFs and EMFs respectively, see figure 3 and below).

The nucleotide sequence information provided herein was obtained by sequencing the WSBV genome using a shotgun sequencing method known in the art. WSBV genomic DNA was initially obtained for sequencing by extraction and purification of viral DNA from infected shrimp tissues using the method of Yang et al. (*J. Virological Methods*, 67:1-4 (1997)), which is hereby incorporated by reference. The nucleotide sequences provided herein are highly accurate, although not necessarily a 100% perfect, representation of the nucleotide sequence of the WSBV genome.

Using the information provided herein together with routine cloning and sequencing methods, one of ordinary skill in the art is able to identify, clone and sequence all "representative fragments" of interest including open reading frames (ORFs) encoding a large variety of WSBV proteins. In very rare instances, this may reveal a nucleotide sequence error present in the nucleotide sequence disclosed herein. Thus, once the present invention is made available (i.e., the information in the Sequence Listing and figures in a useable form), resolving a rare sequencing error would be well within the skill of the art. Nucleotide sequence editing software is publicly available.

Even if all of the very rare sequencing errors in the sequences herein disclosed were corrected, the resulting nucleotide sequence would still be at least 90% identical, and more likely 99% identical, and most likely 99.99% identical to the nucleotide sequence provided herein.

Thus, the present invention further provides nucleotide sequences that are at least 90% identical, or greater, to the nucleotide sequences of the present invention in a form which can be readily used, analyzed and interpreted by the skilled artisan. Methods for determining whether a nucleotide sequence is at least 90% identical to the nucleotide sequence of the present invention are routine and readily available to the skilled artisan. For example, the well known BLAST algorithm can be used to generate the percent identity of nucleotide sequences.

The present invention further provides a prediction of all of the genes within the WSBV genome. This information is provided in Figure 3. The information in the figures can be used to generate WSBV detection kits, antiviral agents, expression arrays, microfluidic devices, individual gene fragments, proteins, antibodies, promoters, protein and nucleotide based viral screens and the like, and to identify commercially important genes and gene products.

#### Specific Embodiments

##### Computer Related Embodiments

The nucleotide sequences provided in the present invention, a representative fragment thereof, or nucleotide sequences at least 90% identical to these sequences, may be "provided" in a variety of mediums to facilitate use thereof. As used herein, "provided" refers to a manufacture, other than an isolated nucleic acid molecule, that contains a nucleotide sequence of the present invention, i.e., the nucleotide sequences provided in the present invention, a representative fragment thereof, or nucleotide sequences at least 90% identical to these sequences. Such a manufacture provides the WSBV genome or a subset thereof (e.g., a WSBV open reading frame (ORF)) in a form that allows a skilled artisan to examine the manufacture using means not directly applicable to examining the WSBV genome or a subset thereof as it exists in nature or in purified form.

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium that can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. One such medium is provided with the present application, namely, the present application contains computer readable medium (CD-R) that has the sequence contigs provided/recorded thereon in ASCII text format in a Sequence Listing.

As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for

recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide or amino acid sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as OB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing the nucleotide sequences of the present invention, a representative fragment thereof, or nucleotide sequences at least 90% identical to these sequences, in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. Software which implements the BLAST (Altschul *et al*, *J. Mol. Biol.* 215:403-410 (1990)) and BLAZE (Brutlag *et al*, *Comp. Chem.* 17:203-207 (1993)) search algorithms on a Sybase system can be used to identify open reading frames (ORFs) within the WSBV genome that contain homology to ORFs or proteins from other organisms. Such ORFs are protein-encoding fragments within the WSBV genome and are useful in producing commercially important proteins such as proteins used as drug or antiviral targets.

The present invention further provides systems, particularly computer-based systems, which contain the sequence information described herein. Such systems are designed to identify commercially important fragments of the WSBV genome.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based system are suitable for use in the present invention. Such system

can be changed into a system of the present invention by utilizing the sequence information provided on the CD-R, or a subset thereof without any experimentation.

As stated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs that are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the WSBV genome which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are available and can be used in the computer-based systems of the present invention. Examples of such software include, but is not limited to, MacPattern (EMBL), BLASTN and BLASTX (NCBIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems.

As used herein, a "target sequence" can be any DNA or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 100 amino acids or from about 30 to 300 nucleotide residues. However, it is well recognized that searches for commercially important fragments of the WSBV genome, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) is chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzymatic active sites and signal sequences. Nucleic acid target motifs include,

but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

A variety of structural formats for the input and output means can be used to input and output the information in the computer-based systems of the present invention. A preferred format for an output means ranks fragments of the WSBV genome possessing varying degrees of homology to the target sequence or target motif. Such presentation provides a skilled artisan with a ranking of sequences which contain various amounts of the target sequence or target motif and identifies the degree of homology contained in the identified fragment.

A variety of comparing means can be used to compare a target sequence or target motif with the data storage means to identify sequence fragments of the WSBV genome. Software which implements the BLAST and BLAZE algorithms (Altschul et al., J Mol. Biol. 215:403-410 (1990)) can be used to identify open reading frames within the WSBV genome. A skilled artisan can readily recognize that any one of the publicly available homology search programs can be used as the search means for the computer-based systems of the present invention.

One application of this embodiment is provided in Figure 1. Figure 1 provides a block diagram of a computer system 102 that can be used to implement the present invention. The computer system 102 includes a processor 106 connected to a bus 104. Also connected to the bus 104 are a main memory 108 (preferably implemented as random access memory, RAM) and a variety of secondary storage devices 110, such as a hard drive 112 and a removable medium storage device 114. The removable medium storage device 114 may represent, for example, a floppy disk drive, a CD-ROM drive, a magnetic tape drive, etc. A removable storage medium 116 (such as a floppy disk, a compact disk, a magnetic tape, etc.) containing control logic and/or data recorded therein may be inserted into the removable medium storage device 114. The computer system 102 includes appropriate software for reading the control logic and/or the data from the removable storage medium 116 once inserted in the removable medium storage device 114.

The nucleotide sequences of the present invention may be stored in a well known manner in the main memory 108, any of the secondary storage devices 110, and/or a removable storage medium 116. Software for accessing and processing the genomic sequence (such as search tools, comparing tools, etc.) reside in main memory 108 during execution.

### Biochemical Embodiments

#### Nucleic Acid Fragments

Another embodiment of the present invention is directed to isolated fragments of the WSBV genome. The fragments of the WSBV genome of the present invention include, but are not limited to, fragments that encode peptides, hereinafter open reading frames (ORFs) and fragments which modulate the expression of an operably linked ORF. Some of these fragments are identified and described in Figure 3. The isolated nucleic acid molecules of the present invention include, but are not limited to, single stranded and double stranded DNA, and single stranded RNA.

"Nucleotide sequence" refers to a heteropolymer of deoxyribonucleotides. Generally, DNA segments encoding the polypeptides and proteins provided by this invention are assembled from fragments of the WSBV genome or single nucleotides, short oligonucleotide linkers, or from a series of oligonucleotides, to provide a synthetic nucleic acid molecule.

As used herein, an "isolated nucleic acid molecule" or an "isolated fragment of the WSBV genome" refers to a nucleic acid molecule possessing a specific nucleotide sequence which has been subjected to purification means to reduce, from the composition, the number of compounds which are normally associated with the composition. A variety of purification means can be used to generate the isolated fragments of the present invention. These include, but are not limited to, methods that separate constituents of a solution based on charge, solubility, or size.

In one embodiment, WSBV DNA can be mechanically sheared to produce fragments of about 2kb, 10kb, or 15-20 kb in length. These fragments can then be used to generate a WSBV library by inserting them into vectors, such as plasmid or lambda vectors, using methods well known in the art. Primers flanking each fragment, for example an ORF, can then be generated using nucleotide sequence information provided in the present invention. PCR cloning can then be used to isolate the ORF from the WSBV DNA library. PCR cloning is well known in the art. Thus, given the availability of the present identified gene coding sequences of the WSBV genome, it is routine experimentation to isolate any ORF, or other fragment of the assembly of the present invention, particularly using the information provided in Figure 3. Such fragments can be applied to an array, microfluidic device, or other detection kit format and used to detect expression of a viral gene (see below).

As used herein, an "open reading frame" (ORF) means a series of triplets coding for amino acids without any termination codons and is a sequence translatable into protein. A

skilled artisan can readily identify ORFs in the WSBV genome using the gene coding sequences provided herein and/or the computer-based systems of the present invention.

As used herein, an "expression modulating fragment" (EMF) means a series of nucleotide molecules which modulates the expression of an operably linked ORF or another EMF.

As used herein, a viral sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of viral EMFs are fragments which induce the expression of an operably linked viral ORF in response to a specific host regulatory factor or physiological event, such as a host anti-viral response.

EMF sequences can be identified within the WSBV genome by their proximity to the ORFs identified using the computer-based systems of the present invention. EMFs may be found immediately 5' to the ORF. Alternatively, EMFs can be identified using known EMFs as a target sequence or target motif in the computer-based systems of the present invention.

The presence and activity of an EMF can be confirmed using an EMF trap vector. An EMF trap vector contains a cloning site 5' to a marker sequence. A marker sequence encodes an identifiable phenotype, such as antibiotic resistance or a complementing nutrition auxotrophic factor, which can be identified or assayed when the EMF trap vector is placed within an appropriate host under appropriate conditions. An EMF will modulate the expression of an operably linked marker sequence. A sequence that is suspected of being an EMF is cloned in all three reading frames in one or more restriction sites upstream from the marker sequence in the EMF trap vector. The vector is then transformed into an appropriate host using known procedures and the phenotype of the transformed host is examined under appropriate conditions.

The sequences falling within the scope of the present invention are not limited to the specific sequences herein described, but also include variations thereof. Variations can be routinely determined by comparing the sequence provided in the present invention, or a representative fragment thereof, with a sequence from another WSBV isolate. Furthermore, to accommodate the degeneracy of the genetic code, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another that encodes the same amino acid is expressly contemplated.



Any specific sequence disclosed herein can be readily screened for errors by resequencing a particular fragment, such as an ORF, in both directions (i.e., sequence both strands). Alternatively, error screening, or variant detection, can be performed by sequencing corresponding polynucleotides of WSBV origin isolated by using part or all of the fragments in question as a probe or primer.

#### Nucleic Acid Fragment Uses

The nucleic acid molecules of the present invention are useful for probes, primers, chemical intermediates, and in biological assays. The nucleic acid molecules are useful as hybridization probes for viral messenger RNA, viral transcript/cDNA, and viral genomic DNA to isolate full-length viral cDNA and viral genomic clones encoding the peptides described in Figure 3, and for use in viral screens and antiviral agents.

Oligonucleotide probes have long been used to detect complementary nucleic acid sequences in a nucleic acid of interest (the "target" nucleic acid) in the form of detection kits/reagents. In some assay formats, the oligonucleotide probe is tethered, i.e., by covalent attachment, to a solid support, and arrays of oligonucleotide probes immobilized on solid supports have been used to detect specific nucleic acid sequences in a target nucleic acid. See, e.g., PCT patent publication Nos. WO 89/10977 and 89/11548. In other formats, the detection reagents are supplied in solution.

The probe can correspond to any sequence along the entire length of the nucleic acid molecules provided in the figures. However, as discussed, fragments are not to be construed as encompassing fragments disclosed prior to the present invention.

Each of the ORFs of the WSBV genome that can be routinely identified using the computer system of the present invention can be used in numerous ways as polynucleotide reagents. The sequences can be used as diagnostic probes or diagnostic amplification primers to detect the expression of a particular gene or groups of genes. This is particularly useful in the form of nucleic acid arrays employing 1 or more, 10 or more, 100 or more, or most to all of the WSBV ORFs in a single array.

The nucleic acid molecules are also useful as primers for PCR to amplify any given region of a nucleic acid molecule and are useful for synthesizing antisense molecules of desired length and sequence.

The nucleic acid molecules are useful as primers for the 5' nuclease PCR assay (hereafter referred to as the TaqMan assay). The TaqMan assay provides a sensitive and rapid means of detecting viral nucleic acid and therefore is well suited for use in viral screening

applications such as detection kits. The TaqMan assay detects the accumulation of a specific amplified product during PCR. The TaqMan assay utilizes an oligonucleotide probe labeled with a fluorescent reporter dye at the 5' end of the probe and a quencher dye at the 3' end of the probe. During the PCR reaction, the 5' nuclease activity of DNA polymerase cleaves the probe, thereby separating the reporter dye and the quencher dye and resulting in increased fluorescence of the reporter. Accumulation of PCR product is detected directly by monitoring the increase in fluorescence of the reporter dye. The 5' nuclease activity of DNA polymerase cleaves the probe between the reporter and the quencher only if the probe hybridizes to the target and is amplified during PCR. Therefore, only the target sequence of interest is detected.

Preferred TaqMan primer and probe sequences are disclosed in Figure 3. It will be apparent to one of skill in the art that the disclosed primers and probes of the present invention are useful as diagnostic probes or amplification primers for screening for the presence of WSBV in a biological sample or for isolating or screening particular WSBV genes.

The nucleic acid molecules are also useful for expressing antigenic portions of the WSBV proteins that can then be used, for example, to develop antibodies to the viral antigens.

The nucleic acid molecules are also useful as hybridization probes for determining the presence, level, form and distribution of WSBV nucleic acid expression. Accordingly, the probes can be used to detect the presence of, or to determine levels of, a specific viral nucleic acid molecule, either DNA or RNA, in cells and tissues of shrimp or other organisms under moderate or stringent conditions. One example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45 °C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65 °C. Examples of moderate to low stringency hybridization conditions are well known in the art. Furthermore, probes corresponding to the viral peptides described herein can be used to assess expression and/or gene copy number in a given infected cell, tissue, or organism. For example, Northern blots can be used for RNA detection, Southern blots can be used for DNA detection, and Western blots can be used for peptide/protein detection. These uses are relevant for detecting the presence of virus in shrimp as well as for monitoring the distribution of virus throughout various cells and tissues of shrimp during the course of viral infection.

In addition, each of the expression modulating fragments (EMFs) can be used in DNA-protein binding assays to screen for modulating peptides which may be present in cells and

tissues of shrimp or other organisms. This use is relevant for obtaining the specific host regulatory factors that interact with promoters in the WSBV genome.

Nucleic acid expression assays are also useful for drug screening to identify compounds that modulate viral nucleic acid expression. The invention thus provides a method for identifying a compound that can be used to treat viral infection. The method typically includes assaying the ability of the compound to modulate the expression of viral nucleic acid and thus identifying a compound that can be used to treat a viral infection. The assays can be performed in cell-based and cell-free systems. Cell-based assays include cells infected with virus particles or recombinant cells genetically engineered to express specific viral nucleic acid sequences. Cell-free assays can be used to detect the ability of a compound to directly bind to a nucleic acid fragment or protein.

The assay for viral nucleic acid expression can involve direct assay of nucleic acid levels, such as mRNA levels, or on collateral compounds involved in the signal pathway. Furthermore, the expression of host cell genes that are up- or down-regulated in response to the viral protein can also be assayed. In this embodiment the regulatory regions of these genes can be operably linked to a reporter gene such as luciferase.

Thus, modulators of viral gene expression can be identified by a method wherein a cell infected with virus is contacted with a candidate compound and the expression of viral mRNA determined. The level of expression of viral mRNA in the presence of the candidate compound is compared to the level of expression of viral mRNA in the absence of the candidate compound. The candidate compound can then be identified as a modulator of viral nucleic acid expression based on this comparison and be used, for example, to disrupt viral replication. When expression of viral mRNA is statistically significantly greater in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of viral nucleic acid expression. When viral nucleic acid expression is statistically significantly less in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of viral nucleic acid expression.

The invention further provides methods of treating viral infection, with the nucleic acid as a target, using a compound identified through drug screening as a gene modulator to modulate viral nucleic acid expression in cells and tissues infected with the virus. Modulation includes both up-regulation (i.e. activation or agonization) or down-regulation (suppression or antagonization) of nucleic acid expression. Generally, viral nucleic acid expression is down-regulated to prevent viral replication and treat viral infection.

Alternatively, a modulator for viral nucleic acid expression can be a small molecule or drug identified using the screening assays described herein as long as the drug or small molecule modulates viral nucleic acid expression in the cells and tissues infected with the virus.

The nucleic acid molecules are also useful for monitoring the effectiveness of modulating compounds on the expression or activity of the viral gene in a treatment regimen. Thus, the gene expression pattern can serve as a barometer for the continuing effectiveness of treatment with the compound, particularly with compounds to which a virus can develop resistance. The gene expression pattern can also serve as a marker indicative of a physiological response of the virus to the compound. Accordingly, such monitoring would allow either increased administration of the compound or the administration of alternative compounds to which the virus has not become resistant.

The nucleic acid molecules are also useful as antisense constructs to control viral gene expression in infected cells and tissues. A DNA antisense nucleic acid molecule is designed to be complementary to, and therefore bind to, a region of the viral gene necessary for transcription, thereby preventing transcription and hence production of viral protein. An antisense RNA or DNA nucleic acid molecule would hybridize to the viral mRNA and thus block translation of viral mRNA into protein by the host cell's translational machinery. Alternatively, a class of antisense molecules can be used to inactivate viral mRNA in order to decrease expression of viral nucleic acid and inhibit viral replication or function. These molecules can therefore be used to treat viral infection. This technique involves cleavage of viral mRNA by ribozymes that recognize one or more regions of viral mRNA that attenuate the ability of the mRNA to be translated by host cell translational machinery. Possible regions include coding or control regions; particularly coding or control regions encoding or regulating proteins that play critical roles in viral function or replication, such as entry into the nucleus of the host cell or virion assembly.

The nucleic acid molecules of the present invention can be employed to create transgenic viral resistant shrimp. Several possible mechanisms could be employed to impart WSBV resistance to shrimp using the nucleic acid and protein coding sequences provided by the present invention. One possible mechanism of imparting WSBV resistance to shrimp involves transforming shrimp cells with viral nucleic acids that express an attenuated virion coat protein such that when the transgenic shrimp is infected with WSBV, the expressed coat protein envelopes the virus and thereby prevents translation of the viral DNA. In this example, the virion coat protein can either be constitutively expressed or regulated by a promoter that is activated upon WSBV infection. Shrimp cells can be transformed with viral DNA under

suitable conditions known in the art. The WSBV construct in a vector can be microinjected directly into host cells using micropipettes, [Crossway, *Mol. Gen. Genetics*, 202:179-85 (1985)], or using polyethylene glycol [Krens *et al.*, *Nature*, 296:72-74 (1982)]. Alternatively, shrimp cells may be transformed by incubating the shrimp cells or tissue with an inoculum of bacteria that have been transformed with a vector comprising a gene that imparts WSBV resistance. The transformed shrimp cells are then grown and regenerated into shrimp such that the proteins expressed by the transformed cells impart WSBV resistance to the shrimp.

#### Nucleic Acid Arrays and Detection Reagents

The present invention further provides detection reagents and kits, such as arrays or microarrays, of nucleic acid molecules that are based on the novel WSBV sequence information provided in the present invention and particularly the transcript information (SEQ ID NOS: 2, 4, 6, . . . 280, 282, 284, 286-293) provided in Figure 3 and polymorphic sites on these transcripts summarized in Table 1.

As used herein "Arrays" or "Microarrays" refers to an array of distinct polynucleotides or oligonucleotides synthesized on a substrate, such as paper, nylon or other type of membrane, filter, chip, glass slide, or any other suitable solid, or semi-solid support. The development of arraying technologies such as photolithographic synthesis of a nucleic acid array and high density spotting of cDNA products has provided methods for making very large arrays of oligonucleotide probes in very small areas. See U.S. Pat. No. 5,143,854 and PCT patent publication Nos. WO 90/15070 and 92/10092. Microfabricated arrays of large numbers of oligonucleotide probes, called "DNA chips", offer great promise for a wide variety of applications. In one embodiment, the microarray is prepared and used according to the methods described in US Patent 5,837,832, Chee *et al.*, PCT application W095/11995 (Chee *et al.*), Lockhart, D. J. *et al.* (1996; *Nat. Biotech.* 14: 1675-1680) and Schena, M. *et al.* (1996; *Proc. Natl. Acad. Sci.* 93: 10614-10619), all of which are incorporated herein in their entirety by reference. In other embodiments, such arrays are produced by the methods described by Brown *et al.*, US Patent No. 5,807,522.

The microarray or detection kit is preferably composed of a large number of unique, single-stranded nucleic acid sequences, usually either synthetic antisense oligonucleotides or fragments of cDNAs, fixed to a solid support. The oligonucleotides are preferably about 6-60 nucleotides in length, more preferably 15-30 nucleotides in length, and most preferably about 20-25 nucleotides in length. For a certain type of microarray or detection kit, it may be

preferable to use oligonucleotides that are only 7-20 nucleotides in length. For others, such as cDNA, longer lengths are possible and preferable. These can be of the order of 1kb or more.

The microarray or detection kit may contain oligonucleotides that cover the known 5' or 3' sequence, sequential oligonucleotides that cover the full-length sequence; or unique oligonucleotides selected from particular areas along the length of the sequence.

Polynucleotides used in the microarray or detection kit may be oligonucleotides that are specific to a viral gene or genes of interest.

In order to produce oligonucleotides to a known sequence for a microarray or detection kit, the viral gene(s) of interest (or an ORF identified from the contigs of the present invention) is typically examined using a computer algorithm which starts at the 5' or at the 3' end of the nucleotide sequence. Typical algorithms will then identify oligomers of defined length that are unique to the gene, have a GC content within a range suitable for hybridization, and lack predicted secondary structure that may interfere with hybridization. In certain situations it may be appropriate to use pairs of oligonucleotides on a microarray or detection kit. The "pairs" will be identical, except for one nucleotide that preferably is located in the center of the sequence. The second oligonucleotide in the pair (mismatched by one) serves as a control. The number of oligonucleotide pairs may range from one to two millions. The oligomers are synthesized at designated areas on a substrate using a light-directed chemical process. The substrate may be paper, nylon or other type of membrane, filter, chip, glass slide or any other suitable solid support.

In another aspect, an oligonucleotide may be synthesized on the surface of the substrate by using a chemical coupling procedure and an ink jet application apparatus, as described in PCT application W095/251116 (Baldeschweiler et al.) which is incorporated herein in its entirety by reference. In another aspect, a "gridded" array analogous to a dot (or slot) blot may be used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical bonding procedure. An array, such as those described above, may be produced by hand or by using available devices (slot blot or dot blot apparatus), materials (any suitable solid support), and machines (including robotic instruments), and may contain 8, 24, 96, 384, 1536, 6144 or more oligonucleotides, or any other number which lends itself to the efficient use of commercially available instrumentation.

In other embodiments, the array or detection reagent/kit can be produced by spotting a nucleic acid molecule onto the surface of a substrate (See Brown et. al., US Patent No.

5,807,522). In such embodiments, PCR primers to one or more nucleic acid fragments are used to generate nucleic acid molecules suitable for deposition onto a substrate.

In order to conduct sample analysis using a microarray or detection kit, viral nucleic acid is isolated from a biological sample infected with WSBV and the viral nucleic acid is made into hybridization probes. Viral nucleic acid may be isolated from biological samples obtained from fluids, cultured cells, biopsies, or other tissue preparations from a shrimp or other organism of interest that is infected with WSBV. Viral mRNA is isolated, and cDNA is produced and used as a template to make antisense RNA (aRNA). The aRNA is amplified in the presence of fluorescent nucleotides, and labeled probes are incubated with the microarray or detection kit so that the probe sequences hybridize to complementary oligonucleotides of the microarray or detection kit. Incubation conditions are adjusted so that hybridization occurs with precise complementary matches or with various degrees of less complementarity. After removal of nonhybridized probes, a scanner is used to determine the levels and patterns of fluorescence. The scanned images are examined to determine degree of complementarity and the relative abundance of each oligonucleotide sequence on the microarray or detection kit. A detection system may be used to measure the absence, presence, and amount of hybridization for all of the distinct viral sequences simultaneously. This data may be used for large scale correlation studies on the sequences, expression patterns, mutations, variants, or polymorphisms among viral isolates.

Using such arrays, the present invention provides methods to identify the expression of one or more of the ORFs of the present invention. In detail, such methods comprise incubating a test sample with one or more nucleic acid molecules and assaying for binding of the nucleic acid molecule with components within the test sample. Such assays will typically involve arrays comprising most, if not all of the genes in the WSBV genome, or rationally selected subsets thereof. The genomic sequence (SEQ ID NO: 1) and transcript sequences (SEQ ID NOS: 2, 4, 6, . . . 280, 282, 284, 286-293) of the WSBV genome of the present invention are provided in Figure 2 and Figure 3 and polymorphic sites on these transcripts summarized in Table 1.

Conditions for incubating a nucleic acid molecule with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid molecule used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or array assay formats can readily be adapted to employ the novel fragments of the WSBV genome disclosed herein. Examples of such assays can be found in Chard, T, *An Introduction to*

*Radioimmunoassay and Related Techniques*, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G. R. et al., *Techniques in Immunocytochemistry*, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., *Practice and Theory of Enzyme Immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology*, Elsevier Science Publishers, Amsterdam, The Netherlands (1985).

The test samples of the present invention include, but are not limited to, nucleic acid extracts, cells, and protein or membrane extracts of cells infected with WSBV. The test sample used in the above-described method will vary based on the assay format, the nature of the detection method, and the tissues, cells, or extracts used as the sample to be assayed. Methods for preparing nucleic acid extracts or for preparing cells are well known in the art and can readily be adapted in order to obtain a sample that is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention.

Specifically, the invention provides a compartmentalized kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the nucleic acid molecules that can bind to a fragment of the WSBV genome disclosed herein; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting the presence of a bound nucleic acid. Preferred kits will include detection reagents/arrays/chips/microfluidic devices that are capable of detecting the expression of 1 or more, 10 or more, 100 or more, or most or all of the genes expressed in WSBV, particularly the genes provided in Figure 3.

In detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers, strips of plastic, glass or paper, or arraying material such as silica. Such containers allow one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers may include a container which will accept the test sample, a container which contains the nucleic acid probe, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound probe. One skilled in the art will readily recognize that the previously unidentified ORFs that can be routinely identified using the sequence information disclosed herein can be



readily incorporated into one of the established kit formats which are well known in the art, particularly expression arrays.

#### Protein/Peptide Molecules

The present invention provides nucleic acid sequences that encode WSBV protein molecules (Figure 3). The peptide sequences provided in Figure 3, as well as the obvious variants described herein, and using the information in Figure 3, will be referred to herein as the WSBV peptides of the present invention or peptides/polypeptides/proteins of the present invention.

Enzymes and other viral proteins are produced during viral activity and replication and can be used as targets for screening and quantitating a particular virus, or as antiviral targets. Like viruses in general, WSBV utilizes the resources of the host cell for production of viral proteins. Viral proteins can be detected using an antibody, or binding portion thereof, to the protein or a probe that recognizes proteins or peptides of the present invention. Viral antigens present either on the surface or within the infected cell can be detected by various antibody tests, including immunofluorescence or enzyme immunoassay (EIA). Virus or antigen released from infected cells can be detected by such antibody tests as enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA), or latex agglutination (LA). Protein-based tests such as these for WSBV antigens are useful for detecting outbreaks of WSBV in shrimp populations.

The present invention provides isolated peptide and protein molecules that comprise, consist essentially of, or consist of the amino acid sequences of the WSBV peptides disclosed in Figure 3, (which are encoded by the transcript sequences that are also shown in Figure 3), as well as all obvious variants of these peptides that are within the art to make and use. Some of these variants are described in detail below.

As used herein, a peptide is said to be "isolated" or "purified" when it is substantially free of cellular material or free of chemical precursors or other chemicals. The peptides of the present invention can be purified to homogeneity or other degrees of purity. The level of purification will be based on the intended use. The critical feature is that the preparation allows for the desired function of the peptide, even if in the presence of considerable amounts of other components.

In some uses, "substantially free of cellular material" includes preparations of the peptide having less than about 30% (by dry weight) other proteins (i.e., contaminating protein), less than about 20% other proteins, less than about 10% other proteins, or less than about 5% other

proteins. When the peptide is recombinantly produced, it can also be substantially free of culture medium, i.e., culture medium represents less than about 20% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of the peptide in which it is separated from chemical precursors or other chemicals that are involved in its synthesis. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of a WSBV peptide having less than about 30% (by dry weight) chemical precursors or other chemicals, less than about 20% chemical precursors or other chemicals, less than about 10% chemical precursors or other chemicals, or less than about 5% chemical precursors or other chemicals.

An isolated WSBV peptide can be purified from cells infected with WSBV, purified from cells that have been altered to express it (recombinant), or synthesized using known protein synthesis methods. For example, a nucleic acid molecule encoding the peptide can be cloned into an expression vector, the expression vector introduced into a host cell, and the protein expressed in the host cell. The protein can then be isolated from the host cells by an appropriate purification scheme using standard protein purification techniques. Many of these techniques are described in detail below.

Accordingly, the present invention provides proteins that consist of the amino acid sequences provided in Figure 3 (SEQ ID NOS:3, 5, 7...281, 283, 285), for example, proteins encoded by the transcript/cDNA nucleic acid sequences also shown in Figure 3 (SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293). A protein consists of an amino acid sequence when the amino acid sequence is the final amino acid sequence of the protein.

The present invention further provides proteins that consist essentially of the amino acid sequences provided in Figure 3 (SEQ ID NOS:3, 5, 7...281, 283, 285), for example, proteins encoded by the transcript/cDNA nucleic acid sequences also shown in Figure 3 (SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293). A protein consists essentially of an amino acid sequence when such an amino acid sequence is present with only a few additional amino acid residues, for example from about 1 to about 100 or so additional residues, typically from 1 to about 20 additional residues in the final protein.

The present invention further provides proteins that comprise the amino acid sequences provided in Figure 3 (SEQ ID NOS:3, 5, 7...281, 283, 285), for example, proteins encoded by the transcript/cDNA nucleic acid sequences also shown in Figure 3 (SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293). A protein comprises an amino acid sequence when the amino acid sequence is at least part of the final amino acid sequence of the protein. In such a fashion, the

protein can be only the peptide or have additional amino acid molecules, such as amino acid residues (contiguous encoded sequence) that are naturally associated with it or heterologous amino acid residues/peptide sequences. Such a protein can have a few additional amino acid residues or can comprise several hundred or more additional amino acids. The preferred classes of proteins that are comprised of the peptides of the present invention are the naturally occurring mature proteins. A brief description of how various types of these proteins can be made/isolated is provided below.

The peptides of the present invention can be attached to heterologous sequences to form chimeric or fusion proteins. Such chimeric and fusion proteins comprise a WSBV peptide operatively linked to a heterologous protein having an amino acid sequence not substantially homologous to the WSBV peptide. "Operatively linked" indicates that the WSBV peptide and the heterologous protein are fused in-frame. The heterologous protein can be fused to the N-terminus or C-terminus of the WSBV peptide.

In some uses, the fusion protein does not affect the activity of the WSBV peptide per se. For example, the fusion protein can include, but is not limited to, enzymatic fusion proteins, for example beta-galactosidase fusions, yeast two-hybrid GAL fusions, poly-His fusions, MYC-tagged, HI-tagged and Ig fusions. Such fusion proteins, particularly poly-His fusions, can facilitate the purification of recombinant WSBV peptide. In certain host cells, expression and/or secretion of a protein can be increased by using a heterologous signal sequence.

A chimeric or fusion protein can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different protein sequences are ligated together in-frame in accordance with conventional techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and re-amplified to generate a chimeric gene sequence (see Ausubel et al., *Current Protocols in Molecular Biology*, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST protein). A WSBV peptide-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the WSBV peptide.

As mentioned above, the present invention also provides and enables obvious variants of the amino acid sequence of the proteins of the present invention, such as naturally occurring mature forms of the peptide, sequence variants of the peptides, non-naturally occurring recombinantly derived variants of the peptides, and orthologs and paralogs of the peptides. Such

variants can readily be generated using art-known techniques in the fields of recombinant nucleic acid technology and protein biochemistry. It is understood, however, that variants exclude any amino acid sequences disclosed prior to the invention.

Such variants can readily be identified/made using molecular techniques and the sequence information disclosed herein. Further, such variants can readily be distinguished from other peptides based on sequence and/or structural homology to the WSBV peptides of the present invention. The degree of homology/identity present will be based primarily on whether the peptide is a functional variant or non-functional variant, the amount of divergence present in the paralog protein family and the evolutionary distance between orthologous viruses.

To determine the percent identity of two amino acid sequences or two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In preferred embodiments, at least 30%, 40%, 50%, 60%, 70%, 80%, or 90% or more of the length of a reference sequence is aligned for comparison purposes. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which are introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity and similarity between two sequences can be accomplished using a mathematical algorithm. (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part 1, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991). In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package, using either a Blossom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14,

12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (Devereux, J., *et al.*, *Nucleic Acids Res.* 12(1):387 (1984)), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Myers and W. Miller (CABIOS, 4:11-17 (1989)) which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against sequence databases to, for example, identify other viruses related to WSBV or functionally related protein sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (*J. Mol. Biol.* 215:403-10 (1990)). BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to the nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to the proteins of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.* (*Nucleic Acids Res.* 25(17):3389-3402 (1997)). When utilizing BLAST and gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. The results of one such analysis are provided in Figure 3.

Full-length pre-processed forms, as well as mature processed forms, of proteins that comprise one of the peptides of the present invention can readily be identified as having complete sequence identity to one of the WSBV peptides of the present invention as well as being encoded by the same viral gene as the WSBV peptide provided herein.

Variants of a WSBV peptide can readily be identified as being a WSBV protein having a high degree of sequence homology/identity (also referred to as "significant sequence homology") to at least a portion of the WSBV peptide as well as being encoded by the same viral gene as the WSBV peptide provided herein. Viral genes can readily be determined based on the WSBV sequence information provided in Figure 3. As used herein, two proteins (or a region of the proteins) have significant homology when the amino acid sequences are typically at least about 70-80%, 80-90%, and more typically at least about 90-95% or more homologous. A significantly homologous amino acid sequence, according to the present

invention, will be encoded by a nucleic acid sequence that will hybridize to a WSBV peptide encoding nucleic acid molecule under stringent conditions as more fully described below.

Paralogs of a WSBV peptide can readily be identified as having some degree of significant sequence homology/identity to at least a portion of the WSBV peptide, as being encoded by a gene from WSBV, and as having similar activity or function. Two proteins will typically be considered paralogs when the amino acid sequences typically share at least about 60% or greater, and more typically at least about 70% or greater homology through a given region or domain. Such paralogs will be encoded by a nucleic acid sequence that will hybridize to a WSBV peptide encoding nucleic acid molecule under moderate to stringent conditions as more fully described below.

Orthologs of a WSBV peptide can readily be identified as having some degree of significant sequence homology/identity to at least a portion of the WSBV peptide as well as being encoded by a gene from another virus. Preferred orthologs will be isolated from viruses of commercial or medical importance for the development of broad-spectrum diagnostic and anti-viral agents. Such orthologs will be encoded by a nucleic acid sequence that will hybridize to a WSBV peptide encoding nucleic acid molecule under moderate to stringent conditions, as more fully described below, depending on the degree of relatedness of the two viruses yielding the proteins.

Non-naturally occurring variants of the WSBV peptides of the present invention can readily be generated using recombinant techniques. Such variants include, but are not limited to, deletions, insertions, and substitutions in the amino acid sequence of the WSBV peptide. For example, one class of substitutions is conserved amino acid substitutions. Such substitutions are those that substitute a given amino acid in a WSBV peptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu, and Ile; interchange of the hydroxyl residues Ser and Thr; exchange of the acidic residues Asp and Glu; substitution between the amide residues Asn and Gln; exchange of the basic residues Lys and Arg; and replacements among the aromatic residues Phe and Tyr. Guidance concerning which amino acid changes are likely to be phenotypically silent are found in Bowie *et al.*, *Science* 247:1306-1310 (1990).

Variant WSBV peptides can be fully functional or can lack function in one or more activities, e.g. ability to bind to host cell receptors or ability to form structural components such as the viral nucleocapsid or outer membrane, etc. Fully functional variants typically contain only conservative variation or variation in non-critical residues or in non-critical regions. Figure 3 provides the results of protein analysis and can be used to identify critical domains/regions.

Functional variants can also contain substitutions of similar amino acids that result in no change or an insignificant change in function. Alternatively, such substitutions may positively or negatively affect function to some degree.

Non-functional variants typically contain one or more non-conservative amino acid substitutions, deletions, insertions, inversions, or truncation or a substitution, insertion, inversion, or deletion in a critical residue or critical region.

Amino acids that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham *et al.*, *Science* 244:1081-1085 (1989)), particularly using the results provided in Figure 3. The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as DNA binding. Sites that are critical for virus/host cell receptor binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith *et al.*, *J. Mol. Biol.* 224:899-904 (1992); de Vos *et al.* *Science* 255:306-312 (1992)).

The present invention further provides fragments of the WSBV peptides, in addition to proteins and peptides that comprise and consist of such fragments, particularly those comprising the residues identified in Figure 3. The fragments to which the invention pertains, however, are not to be construed as encompassing fragments that may be disclosed publicly prior to the present invention.

As used herein, a fragment comprises at least 8, 10, 12, 14, 16, or more contiguous amino acid residues from a WSBV peptide. Such fragments can be chosen based on the ability to retain one or more of the biological activities of the WSBV peptide or could be chosen for the ability to perform a function, e.g. bind a substrate or act as an immunogen. Particularly important fragments are biologically active fragments, peptides that are, for example, about 8 or more amino acids in length. Such fragments will typically comprise a domain or motif of the WSBV peptide, e.g., active site or a substrate-binding domain. Further, possible fragments include, but are not limited to, domain or motif containing fragments, soluble peptide fragments, and fragments containing immunogenic structures. Predicted domains and functional sites are readily identifiable by computer programs well known and readily available to those of skill in the art (e.g., PROSITE analysis).

Polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally occurring amino acids. Further, many amino acids, including the terminal amino acids, may be modified by natural processes during the course of viral infection, such as processing and other post-translational modifications by the host cell, or by chemical

modification techniques well known in the art. Common modifications that occur naturally are described in basic texts, detailed monographs, and the research literature, and they are well known to those of skill in the art.

Examples of known modifications include, but are not limited to, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent crosslinks, formation of cystine, formation of pyroglutamate, formylation, gamma carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well known to those of skill in the art and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as *Proteins - Structure and Molecular Properties*, 2nd Ed., T.E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as by Wold, F., *Posttranslational Covalent Modification of Proteins*, B.C. Johnson, Ed., Academic Press, New York 1-12 (1983); Seifter *et al.* (*Meth. Enzymol.* 182: 626-646 (1990)) and Rattan *et al.* (*Ann. N.Y. Acad. Sci.* 663:48-62 (1992)).

Peptides or protein encoding sequences of the present invention can be modified or mutated, either naturally, such as by host cell mechanisms, or by techniques known to those of skill in the art, to disrupt protein formation or protein function, and thereby disrupt viral replication and function. These methods can be used to prevent and/or treat viral infection.

Accordingly, the WSBV peptides of the present invention also encompass derivatives or analogs in which a substituted amino acid residue is not one encoded by the genetic code, in which a substituent group is included, in which the mature WSBV peptide is fused with another compound, such as a compound to increase or decrease the half-life of the WSBV peptide (for example, polyethylene glycol), in which the additional amino acids are fused to the mature WSBV peptide, such as a leader or secretory sequence or a sequence for purification of the mature WSBV peptide or a pro-protein sequence, or in which the WSBV peptide has been modified or mutated, either naturally or recombinantly, to disrupt protein function, and thereby disrupt WSBV function and/or replication.



### Protein/Peptide Uses

The proteins of the present invention can be used in substantial and specific assays related to the functional information provided in the figures; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its binding partner or ligand) in biological samples; and as markers for infected samples in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of viral infection). Where the protein binds or potentially binds to another protein or ligand (such as, for example, a host cell receptor protein), the protein can be used to identify the binding partner/ligand so as to develop a system to identify inhibitors of the binding interaction. Any or all of these uses are capable of being developed into reagent grade or kit format for commercialization as commercial products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

The potential uses of the viral peptides of the present invention are based primarily on the function of the protein. For example, isolated WSBV peptides serve as targets for identifying antiviral agents, particularly for identifying antiviral agents that interfere with viral replication in a host cell infected with a virus that expresses the peptide. Specific and substantial uses for the molecules of the present invention are provided herein. Further uses can readily be determined using the information provided herein, that which is known in the art, and routine experimentation.

The proteins of the present invention (including variants and fragments that may have been disclosed prior to the present invention) are useful for biological assays for viruses that are related to WSBV. Such assays involve any of the known protein functions or activities or properties useful for diagnosis of WSBV infection.

The proteins of the present invention are also useful in virus screening assays, in cell-based or cell-free systems. Cell-based systems can be native, i.e., host cells infected with the virus, as a biopsy or expanded in cell culture. In an alternate embodiment, cell-based assays involve recombinant host cells expressing the viral protein. Cell-based or cell-free systems can be used in assays for protein activity, such as enzymatic activity. Cell-free assays can be used to

detect the ability of a compound to directly bind to a protein or nucleic acid fragment of the present invention.

The polypeptides can be used to identify compounds that modulate activity of the protein. Both the WSBV peptides of the present invention and appropriate variants and fragments can be used in high-throughput screens to assay candidate compounds for the ability to bind to the WSBV peptide. These compounds can be further screened against a functional WSBV peptide to determine the effect of the compound on the WSBV peptide activity. Further, these compounds can be tested in shrimp to determine activity/effectiveness. Compounds can be identified that inactivate the WSBV peptide to a desired degree (antagonists).

Further, the proteins of the present invention can be used to screen a compound for the ability to stimulate or inhibit interaction between the WSBV protein and a target molecule that normally interacts with the WSBV protein, e.g. a host cell receptor. Such assays typically include the steps of combining the WSBV protein with a candidate compound under conditions that allow the WSBV protein, or fragment thereof, to interact with the target molecule, and detecting the formation of a complex between the WSBV protein and the target or detecting the biochemical consequence of the interaction between the WSBV protein and the target, such as any of the associated effects of host cell signal transduction such as protein phosphorylation, cAMP turnover, or adenylate cyclase activation, etc.

Candidate compounds include, for example, 1) peptides such as soluble peptides, including Ig-tailed fusion peptides and members of random peptide libraries (see, e.g., Lam *et al.*, *Nature* 354:82-84 (1991); Houghten *et al.*, *Nature* 354:84-86 (1991)) and combinatorial chemistry-derived molecular libraries made of D- and/or L- configuration amino acids; 2) phosphopeptides (e.g., members of random and partially degenerate, directed phosphopeptide libraries, see, e.g., Songyang *et al.*, *Cell* 72:767-778 (1993)); 3) antibodies (e.g., polyclonal, monoclonal, anti-idiotypic, chimeric, and single chain antibodies as well as Fab, F(ab')<sub>2</sub>, Fab expression library fragments, and epitope-binding fragments of antibodies); and 4) small organic and inorganic molecules (e.g., molecules obtained from combinatorial and natural product libraries).

One candidate compound is a non-virulent soluble fragment of the WSBV peptide that competes for substrate binding, such as for binding to shrimp cellular receptors. Other candidate compounds include non-virulent mutant WSBV peptides or appropriate fragments containing mutations that prevent WSBV virulence and thus compete for substrate. Accordingly, a fragment that competes for substrate, for example with a higher affinity, or a fragment that binds substrate but is inactive or non-virulent, is encompassed by the present invention.

The invention further includes other end point assays to identify compounds that inhibit WSBV activity. The assays typically involve an assay of events in the shrimp cell signal transduction pathway that indicate viral activity. Thus, the phosphorylation of a substrate, activation of a protein, a change in the expression of genes that are up- or down-regulated in a host cell in response to WSBV infection can be assayed.

Any of the viral functions mediated by a WSBV protein can be used as an endpoint assay. These include all of the biochemical or biological events described herein, and in the references cited herein, incorporated by reference for these endpoint assay targets, and other functions known to those of ordinary skill in the art or that can be readily identified using the information provided in the figures, particularly Figure 3.

The proteins of the present invention are also useful in competition binding assays in methods designed to discover compounds that interact with the viral protein (e.g. binding partners and/or ligands). Thus, a compound is exposed to a viral polypeptide under conditions that allow the compound to bind or to otherwise interact with the polypeptide. Soluble viral polypeptide is also added to the mixture. If the test compound interacts with the soluble viral polypeptide, it decreases the amount of complex formed or activity from the viral protein target. This type of assay is particularly useful in cases in which compounds are sought that interact with specific regions of the viral protein. Thus, the soluble polypeptide that competes with the target viral protein region is designed to contain peptide sequences corresponding to the region of interest. See Hodgson, *Bio/technology*, 1992, Sept 10(9), 973-80, for a review of competition binding assays and other receptor screening assays.

To perform cell free drug screening assays, it is sometimes desirable to immobilize either the viral protein, or fragment, or its target molecule to facilitate separation of complexes from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay.

Techniques for immobilizing viral proteins on matrices can be used in the drug screening assays. In one embodiment, a fusion protein can be provided which adds a domain that allows the viral protein to be bound to a matrix. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the cell lysates (e.g., <sup>35</sup>S-labeled) and a candidate drug compound, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads are washed to remove any unbound label, and the matrix immobilized and radiolabel determined directly, or in the supernatant after the complexes are dissociated. Alternatively, the

complexes can be dissociated from the matrix, separated by SDS-PAGE, and the level of a viral protein target ligand, such as a host cell receptor protein, found in the bead fraction quantitated from the gel using standard electrophoretic techniques. For example, either the viral protein or its target ligand can be immobilized utilizing conjugation of biotin and streptavidin using techniques well known in the art. Alternatively, antibodies reactive with the viral protein but which do not interfere with binding of the viral protein to its target ligand can be derivatized to the wells of the plate, and the viral protein trapped in the wells by antibody conjugation. Preparations of a viral protein target ligand and a candidate compound are incubated in the viral protein-presenting wells and the amount of complex trapped in the well can be quantitated. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the viral protein target ligand, or which are reactive with viral protein and compete with the target ligand, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target ligand.

Agents that modulate one of the viral proteins of the present invention can be identified using one or more of the above assays, alone or in combination. It is generally preferable to use a cell-based or cell free system first and then confirm activity in a shrimp, or other organism, infected with WSBV.

Modulators of viral protein activity identified according to these drug screening assays can be used to treat shrimp infected with WSBV. These methods of treatment include the steps of administering a modulator of viral protein activity in a pharmaceutical composition to an organism, such as a shrimp, that is infected with WSBV, the modulator being identified as described herein.

In yet another aspect of the invention, the WSBV proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins which bind to or interact with the viral protein and are involved in viral protein activity, and therefore are targets for inhibiting viral protein activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a viral protein is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g.,

GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, in vivo, forming a viral protein-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the viral protein.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model, such as a shrimp infected with WSBV. For example, an agent identified as described herein (e.g., a viral protein-modulating agent, an antisense viral nucleic acid molecule, a viral protein-specific antibody, or a viral protein-binding partner) can be used in a shrimp, or other organism, infected with WSBV to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal or other model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

The viral proteins of the present invention are also useful for providing targets for diagnosing viral infection. Accordingly, the invention provides methods for detecting the presence, or levels of, the viral protein (or encoding nucleic acid) in an infected cell, tissue, or organism. The method involves contacting a biological sample with a compound capable of interacting with the viral protein such that the interaction can be detected. Such an assay can be provided in a single detection format or a multi-detection format such as an antibody chip array.

One agent for detecting a protein in a sample is an antibody capable of selectively binding to a WSBV protein. A biological sample includes tissues, cells and biological fluids isolated from a shrimp or other infected organism, as well as tissues, cells and fluids present within the infected organism.

In vitro techniques for detection of viral peptide include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence using a detection reagent, such as an antibody or protein binding agent. Alternatively, the peptide can

be detected in vivo in an infected organism by introducing into the subject a labeled anti-peptide antibody or other type of detection agent. For example, the antibody can be labeled with a radioactive marker whose presence and location in an infected organism can be detected by standard imaging techniques. Particularly useful are methods that detect fragments of a peptide in a sample.

### Antibodies

The invention also provides antibodies that selectively bind to one of the WSBV peptides of the present invention, a protein comprising such a peptide, as well as variants and fragments thereof. As used herein, an antibody selectively binds a target peptide when it binds the target peptide and does not significantly bind to unrelated proteins. An antibody is still considered to selectively bind a peptide even if it also binds to other proteins that are not substantially homologous with the target peptide so long as such proteins share homology with a fragment or domain of the peptide target of the antibody. In this case, it would be understood that antibody binding to the peptide is still selective despite some degree of cross-reactivity.

As used herein, an antibody is defined in terms consistent with that recognized within the art: they are multi-subunit proteins produced by a mammalian organism in response to an antigen challenge. The antibodies of the present invention include polyclonal antibodies and monoclonal antibodies, as well as fragments of such antibodies, including, but not limited to, Fab or F(ab')<sub>2</sub>, and Fv fragments.

Many methods are known for generating and/or identifying antibodies to a given target peptide. Several such methods are described by Harlow, *Antibodies*, Cold Spring Harbor Press, (1989). In general, to generate antibodies, an isolated peptide is used as an immunogen and is administered to a mammalian organism, such as a rat, rabbit or mouse. The antibodies generated by the organism in response to the immunogen are then isolated. The full-length protein, an antigenic peptide fragment or a fusion protein can be used. Particularly important fragments are those covering functional domains, such as the domains identified in Figure 3, and domains of sequence homology or divergence between WSBV and other viruses, such as those that can readily be identified using protein alignment methods and as presented in the figures.

Monoclonal antibodies can be produced by hybridomas, which are immortalized cell lines capable of secreting a specific monoclonal antibody. The immortalized cell lines can be created in vitro by fusing two different cell types, usually lymphocytes, one of which is a tumor cell.

The antibodies can be used to isolate one of the proteins of the present invention by standard techniques, such as affinity chromatography or immunoprecipitation. The antibodies can facilitate the purification of the natural viral protein from host cells infected with WSBV, as well as recombinantly produced protein. In addition, such antibodies are useful for detecting the presence of the viral proteins of the present invention in cells or tissues in order to determine the pattern of viral infection among various cells or tissues in a shrimp or other organism over the course of viral infection. Further, such antibodies can be used to detect protein in situ, in vitro, or in a cell lysate or supernatant in order to evaluate the abundance and pattern of viral infection.

The antibodies can also be used to assess subcellular localization of virus particles in host cells. The diagnostic uses can be applied, not only in genetic testing, but also in monitoring a treatment modality. Accordingly, where treatment is ultimately aimed at preventing or halting expression of the WSBV protein, antibodies directed against the protein or relevant fragments can be used to monitor therapeutic efficacy.

The antibodies are also useful diagnostic tools, such as for use as immunological markers for aberrant viral protein analyzed by electrophoretic mobility, isoelectric point, tryptic peptide digest, and other physical assays known to those in the art.

The antibodies are also useful for inhibiting protein function. For example, antibodies may bind directly to viral peptides to block binding of the viral peptide to a binding partner such as a host cell receptor. Antibodies can thereby serve as antiviral agents. Antibodies can be prepared against specific fragments containing sites required for protein function or against intact viral protein that is associated with virulence.

The invention also encompasses kits for using antibodies to detect the presence of a WSBV protein in a biological sample, such as a shrimp cell sample. The kit can comprise antibodies such as a labeled or labelable antibody and a compound or agent for detecting viral protein in a biological sample; means for determining the amount of protein in the sample; means for comparing the amount of protein in the sample with a standard; and instructions for use. Such a kit can be supplied to detect a single protein or epitope or can be configured to detect one of a multitude of epitopes, such as in an antibody detection array. Arrays are described in detail above for nucleic acid arrays and similar methods have been developed for antibody arrays.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the above-described modes for carrying out the invention which are obvious to those skilled in the field of molecular biology or related fields are intended to be within the scope of the following claim.

Certain aspects of the present invention are described in greater detail in the non-limiting examples that follow.



### Examples:

#### Infected prawn

Dead and moribund *P. japonicus* with evident white spots on the inside surface of the crust were collected from a prawn farm and kept at 4 °C.

#### Isolation of nucleocapsids

Hepatopancreata gill and intestine were removed from *P. japonicus* and placed in an ice-bathed beaker, homogenized as a 10% suspension in TESP buffer (50 mmol/l, Tris-HCl, pH 8.5, 10 PMSF), then centrifuged at 6500 x g for 10 min at 4 °C. The supernatant was recentrifuged at 30,000 x g for 30 min at 4 °C. The pellet was suspended in an approximate two volume of TESP buffer containing 1% (v/v) Triton X-100. After centrifugation at 5000 x g for 10 min, the supernatant was centrifuged again at 25,000 x g for 20 min. The pellet was suspended in TESP buffer and differential centrifugation was repeated, and then the precipitate was resuspended in TMP (100 mmol/l, Tris-HCl, pH 7.5, 10 mmol/l MgCl<sub>2</sub>/l, 1 mmol/l PMSF) buffer and treated with DNase and RNase. The mixture was incubated at 37 °C for 15 min and 30 ml of TESP buffer was added. Differential centrifugation was repeated again and the pellet was resuspended in 1 ml of TESP buffer, 1 µl of suspension was dropped on a copper grid, negatively stained with 2% (w/v) uranyl acetate, pH 7.6, and observed using a JEM-100CX II transmission electron microscope.

#### Purification of viral DNA

The nucleocapsid suspension was lysed with 2 ml of GTE buffer (6 mol/l guanidine hydrochloride; 50 mmol/l Tris-HCl; 10 mmol/l EDTA; pH 7.0), slightly mixed, and then centrifuged at 25,000 x g for 10 min at 4 °C. The supernatant was collected and 0.02 vol. of 1 mol/l MgCl<sub>2</sub> and 0.6 vol. of isopropanol were added. After centrifugation, the pellet was picked out with pipette tip, washed twice with 70% ethanol and then dissolved in 1 ml of TE buffer containing 0.5% (w/v) SDS and 0.5 mg/ml proteinase K and incubated at 55 °C for 2-3 hrs. The DNA was precipitated again with 0.01 vol. of 1 mol/l MgCl<sub>2</sub> and 0.25 vol. of isopropanol and dissolved in 0.1 x TE buffer. The viral DNA obtained was quantified by a spectrophotometer.

#### WSBV genomic DNA library construction

#### Construction of a random "shotgun" library:

WSBV genomic DNA was sheared with sonication. Mung Bean nuclease was used to blunt the end. The DNA fragments between 1.8-2kb were recovered from an agarose gel following electrophoresis. The blunt end DNA was cloned into pUC18 vectors. The vector was subsequently transformed into DH5 $\alpha$  cells and plated onto LB plate.

#### Construction of a restriction fragment library:

WSBV genomic DNA was partially digested with Sau3A1 restriction enzymes. DNA fragments between 5-10kb were recovered from the agarose gel. pBluescript vectors were digested by the restriction enzyme and the ends were dephosphorylated. The fragments were cloned into pBluescript vector and transformed into XL-blue competent cells. Subsequently, the DNA plasmid was prepared.

#### Large scale DNA sequencing

##### PCR reactions:

PCR reactions were carried out in a 25  $\mu$ l volume containing 0.2mM dNTP, 1.5mM MgCl<sub>2</sub>, 5  $\mu$ M of each primers, 2.5 unit of Taq polymerase, and a single white colony as template. Take out 1 $\mu$ l as glycerol stocks. PCR reactions were done in a PE 9700. The cycling profile consisted of an initial denaturation at 95°C for 12 min (one cycle) followed by 30 cycles of denaturation at 95°C for 15s, annealing at 58°C for 20s, and extension at 72°C for 2 min. Then 5  $\mu$ l of PCR products were visualized on 1% agarose gels stained with ethidium bromide. Excess primers and dNTPs were removed by digesting the PCR products with exonuclease I and shrimp alkaline phosphatase.

##### DNA sequencing:

Sequencing reactions were carried out in 5 $\mu$ l volume containing 1 $\mu$ l of "BigDye" premix (PE Applied Biosystems), 3.2 $\mu$ M sequencing primer and 30-90ng PCR products. Sequencing reactions were done in a PE 9700. The cycling profile consisted of 30 cycles of denaturation at 95°C for 30s, annealing at 50°C for 30s, and extension at 60°C for 4 min. Excess Dye terminators were removed with ethanol precipitation, and sequencing was carried out on ABI 377 automated sequencer.

##### Sequence analysis

A total of 5795 sequences were assembled in a UNIX system using InnerPeace software designed based on the "Phred, Phrap and Consed" program originally developed by the University of Washington. Sequences were edited and finished as follows: a) for bad

sequence quality, sequencing was repeated; b) for regions with repetitive sequences, which may cause misassembly, primers were designed for walking on the original PCR products or on plasmid DNA; c) for mapping gaps, clones were sequenced that cover the gap; d) for physical gaps, PCR primers were designed between the gaps, then the PCR products were sequenced that cover the gap; e) for gaps that can't be covered by PCR methods, walking on WSBV genomic DNA was applied.

WSBV cDNA library construction and WSBV cDNA clone selection

Poly(A)-mRNA was purified by using the "PolyATtract System1000" kit (Promega). Double stranded cDNA was synthesized and cloned using the "SUPERSCRIP<sup>TM</sup> Plasmid System for cDNA Synthesis and Plasmid Cloning" kit (GIBCO BRL). cDNA clones were transformed into DH10 $\alpha$  cells and then plated. WSBV cDNA clones were selected by DNA hybridization using Dig labeled WSBV genomic DNA as a probe (Dig labeling kit, Boehringer Mannheim). Finally, the plasmid DNA was prepared for automatic sequencing.

### Claims

That which is claimed is:

1. An isolated peptide comprising an amino acid sequence selected from the group consisting of:
  - (a) an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285;
  - (b) a variant of an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285, wherein said variant is encoded by a nucleic acid molecule that hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293;
  - (c) an amino acid sequence of an ortholog of an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285, wherein said ortholog is encoded by a nucleic acid molecule that hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293; and
  - (d) a fragment of an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285, wherein said fragment comprises at least 10 contiguous amino acids.
2. An isolated peptide consisting of an amino acid sequence selected from the group consisting of:
  - (a) an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285.
  - (b) a variant of an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285, wherein said variant is encoded by a nucleic acid molecule that hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293;
  - (c) an amino acid sequence of an ortholog of an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285, wherein said ortholog is encoded by a nucleic acid molecule that hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293; and
  - (d) a fragment of an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285, wherein said fragment comprises at least 10 contiguous amino acids.
3. An isolated antibody that selectively binds to a peptide of claim 1.

4. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence that encodes an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285;

(b) a nucleotide sequence that encodes a variant of an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285, wherein said nucleotide sequence hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293;

(c) a nucleotide sequence that encodes an ortholog of an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285, wherein said nucleotide sequence hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293;

(d) a nucleotide sequence that encodes a fragment of an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285, wherein said fragment comprises at least 10 contiguous amino acids; and

(e) a nucleotide sequence that is the complement of a nucleotide sequence of (a)-(d).

5. An isolated nucleic acid molecule consisting of a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence that encodes an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285;

(b) a variant of an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285, wherein said nucleotide sequence hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293;

(c) a nucleotide sequence that encodes an ortholog of an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285, wherein said nucleotide sequence hybridizes under stringent conditions to the opposite strand of a nucleic acid molecule shown in SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293;

(d) a nucleotide sequence that encodes a fragment of an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285, wherein said fragment comprises at least 10 contiguous amino acids; and

(e) a nucleotide sequence that is the complement of a nucleotide sequence of (a)-(d).

6. A gene chip comprising a nucleic acid molecule of claims 4 or 5.

7. A transgenic non-human organism comprising a nucleic acid molecule of claims 4 or 5.
8. A nucleic acid vector comprising a nucleic acid molecule of claims 4 or 5.
9. A host cell containing the vector of claim 8.
10. A method for producing any of the peptides of claim 1 comprising introducing a nucleotide sequence encoding any of the amino acid sequences in (a)-(d) into a host cell, and culturing the host cell under conditions in which the peptides are expressed from the nucleotide sequence.
11. A method for producing any of the peptides of claim 2 comprising introducing a nucleotide sequence encoding any of the amino acid sequences in (a)-(d) into a host cell, and culturing the host cell under conditions in which the peptides are expressed from the nucleotide sequence.
12. A method for detecting the presence of any of the peptides of claims 1 or 2 in a sample, said method comprising contacting said sample with a detection agent that specifically allows detection of the presence of the peptide in the sample and then detecting the presence of the peptide.
13. A method for detecting the presence of a nucleic acid molecule of claims 4 or 5 in a sample, said method comprising contacting the sample with an oligonucleotide that hybridizes to said nucleic acid molecule under stringent conditions and determining whether the oligonucleotide binds to said nucleic acid molecule in the sample.
14. A method for identifying a modulator of a peptide of claims 1 or 2, said method comprising contacting said peptide with an agent and determining if said agent has modulated the function or activity of said peptide.
15. The method of claim 14, wherein said agent is administered to a host cell comprising an expression vector that expresses said peptide.

16. A method for identifying an agent that binds to any of the peptides of claims 1 or 2, said method comprising contacting the peptide with an agent and assaying the contacted mixture to determine whether a complex is formed with the agent bound to the peptide.
17. A pharmaceutical composition comprising an agent identified by the method of claim 16 and a pharmaceutically acceptable carrier therefor.
18. A method for treating WSBV infection, said method comprising administering to an organism a pharmaceutically effective amount of an agent identified by the method of claim 16.
19. A method for identifying a modulator of the expression of a peptide of claims 1 or 2, said method comprising contacting a cell expressing said peptide with an agent, and determining if said agent has modulated the expression of said peptide.
20. An isolated WSBV peptide having an amino acid sequence that shares at least 70% homology with an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285.
21. A peptide according to claim 20 that shares at least 90 percent homology with an amino acid sequence shown in SEQ ID NOS:3, 5, 7...281, 283, 285.
22. An isolated nucleic acid molecule encoding a WSBV peptide, said nucleic acid molecule sharing at least 80 percent homology with a nucleic acid molecule shown in SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293.
23. A nucleic acid molecule according to claim 22 that shares at least 90 percent homology with a nucleic acid molecule shown in SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293.
24. An isolated nucleic acid detection reagent that is capable of detecting the presence of 1 or more genes from WSBV, wherein said genes are selected from the group consisting of SEQ ID NOS:2, 4, 6...282, 284, 286, 288-292 and 293.
25. The detection reagent of claim 24, wherein said reagent is a nucleic acid array.

26. The array of claim 25, wherein said array is comprised of short oligonucleotides from about 5 to about 100 nucleotides in length.

27. The array of claim 25, wherein said array is comprised of polynucleotides based on the transcript sequences (SEQ ID NOS: 2, 4, 6...280, 282, 284, 286-293), wherein said polynucleotides are from about 100 to about 1000 nucleotides in length.

28. An isolated nucleic acid detection reagent that is capable of detecting the presence of 10 or more genes from WSBV, wherein said genes are selected from the group consisting of SEQ ID NOS:2, 4, 6...282, 284, 286, 288-292 and 293.

29. The detection reagent of claim 28, wherein said reagent is a nucleic acid array.

30. The array of claim 29, wherein said array is comprised of short oligonucleotides from about 5 to about 100 nucleotides in length.

31. The array of claim 29, wherein said array is comprised of polynucleotides based on the transcript sequences (SEQ ID NOS: 2, 4, 6...280, 282, 284, 286-293), wherein said polynucleotides are from about 100 to about 1000 nucleotides in length.

32. An isolated nucleic acid detection reagent that is capable of detecting the presence of 100 or more genes from WSBV, wherein said genes are selected from the group consisting of SEQ ID NOS: 2, 4, 6...282, 284, 286, 288-292 and 293.

33. The detection reagent of claim 32, wherein said reagent is a nucleic acid array.

34. The array of claim 33, wherein said array is comprised of short oligonucleotides from about 5 to about 100 nucleotides in length.

35. The array of claim 33, wherein said array is comprised of polynucleotides based on the transcript sequences (SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293), wherein said polynucleotides are from about 100 to about 1000 nucleotides in length.



36. An isolated nucleic acid detection reagent that is capable of detecting the presence of all genes from WSBV, wherein said genes are selected from the group consisting of SEQ ID NOS:2, 4, 6...282, 284, 286, 288-292 and 293.

37. The detection reagent of claim 36, wherein said reagent is a nucleic acid array.

38. The array of claim 37, wherein said array is comprised of short oligonucleotides from about 5 to about 100 nucleotides in length.

39. The array of claim 38, wherein said array is comprised of polynucleotides based on the transcript sequences (SEQ ID NOS:2, 4, 6...280, 282, 284, 286-293), wherein said polynucleotides are from about 100 to about 1000 nucleotides in length.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
31 May 2001 (31.05.2001)

PCT

(10) International Publication Number  
**WO 01/38351 A2**

(51) International Patent Classification<sup>7</sup>: **C07K 14/00**

(21) International Application Number: PCT/US00/28888

(22) International Filing Date:  
8 November 2000 (08.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
99 1 24717.5 24 November 1999 (24.11.1999) CN

(71) Applicants (for all designated States except US): **PE CORPORATION (NY)** [US/US]; 761 Main Avenue, Norwalk, CT 06859 (US). **THE THIRD INSTITUTE OF OCEANOGRAPHY, STATE OCEANIC ADMINISTRATION, CHINA** [CN/CN]; Haisheng Zhang, Daxue Road #178, Xiamen 361005 (CN). **SINOGENOMAX CO. LTD.** [CN/CN]; Boqin Qiang, 53-3-302, Yuxin, Haidian, Beijing 100096 (CN).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **XU, Xun** [CN/CN];

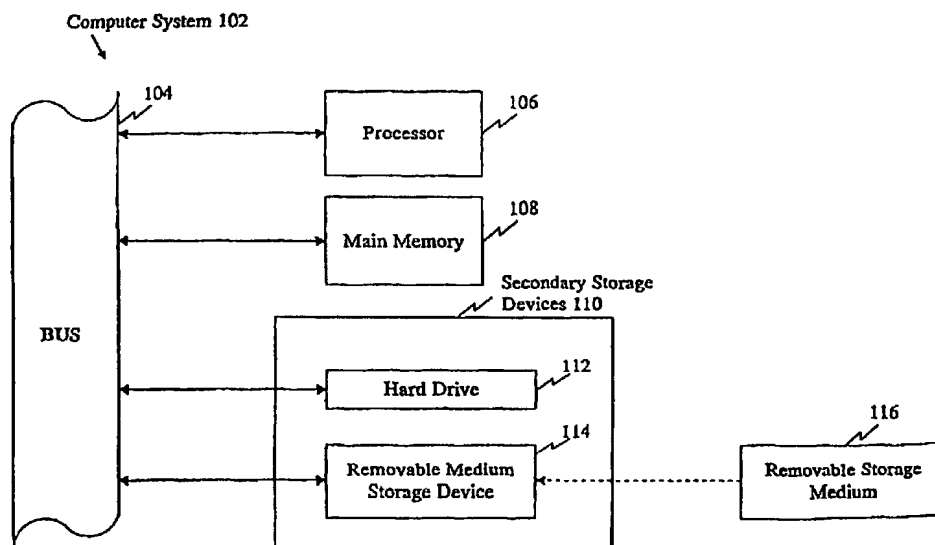
The Third Institute of Oceanography, State Oceanic Administration, China, Daxue Road #178, Xiamen 361005 (CN). **YANG, Feng** [CN/CN]; The Third Institute of Oceanography, State oceanic Administration, China, Daxue Road #178, Xiamen 361005 (CN). **HE, Jun** [CN/CN]; The Third Institute of Oceanography, State Oceanic Administration, China, Daxue Road #178, Xiamen 361005 (CN). **PHAM, Lin-Zuo** [US/US]; 14131 Half Moon Bay Drive, Del Mar, CA 92014 (US). **HE, Mei** [US/CN]; Shanghai GeneCore BioTechnologies, Third Floor, No. 200, Tian Lin Road, Xuhui District, Shanghai 200233 (CN). **YE, Yun** [CN/CN]; Shanghai GeneCore BioTechnologies, Third Floor, No. 200, Tian Lin Road, Xuhui District, Shanghai 200233 (CN). **SHEN, Yan** [CN/CN]; SinoGenoMax Co. Ltd., 53-3-302, Yuxin, Haidian, Beijing 100096 (CN). **KODIRA, Chinnappa** [IN/US]; 20273 Waters Row Terrace, Germantown, MD 20874 (US).

(74) Agent: **MILLMAN, Robert, A.**; Celera Genomics Corp., 45 West Gude Drive C2-4, Rockville, MD 20850 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,

[Continued on next page]

(54) Title: PRIMARY NUCLEOTIDE SEQUENCE OF THE SHRIMP WHITE SPOT BACILLIFORM VIRUS (WSBV), DISCOVERY SYSTEMS CONTAINING THIS SEQUENCE AND DETECTION KITS AND ANTIVIRAL TARGETS FOR DETECTION AND CONTROLLING SHRIMP VIRUS OUTBREAK AND SPREAD



(57) Abstract: The present invention is based on the sequencing and assembly of the WSBV genome. The present invention provides the complete primary nucleotide sequence of the WSBV genome in a series of genomic and predicted transcript sequences. This information is provided in the form of sequences, annotation information, and computer-based systems, and can be used to generate antiviral agents and nucleic acid and protein-based viral detection reagents and kits such as nucleic acid arrays.

WO 01/38351 A2



HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

- (84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— Without international search report and to be republished upon receipt of that report.

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

09/914464

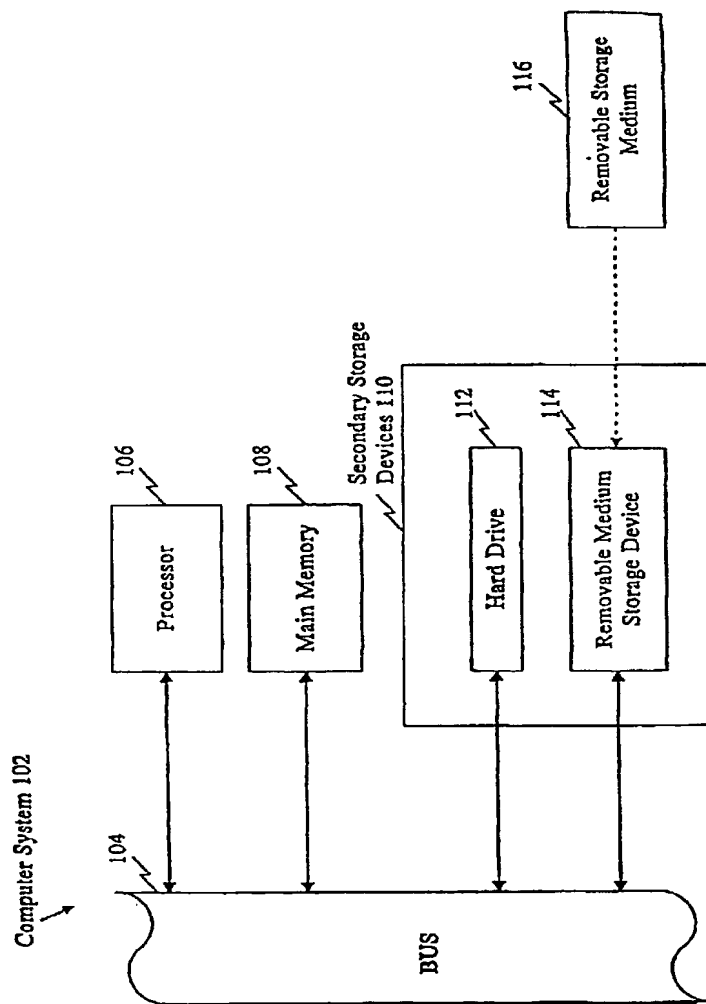


FIGURE 1

FIGURE 2, Sheet 1 of 40

FIGURE 2, Sheet 2 of 40







FIGURE 2. Sheet 5 of 40



FIGURE 2, Sheet 7 of 40

FIGURE 2, Sheet 8 of 40

FIGURE 2. Sheet 9 of 40





FIGURE 2, Sheet 11 of 40





FIGURE 2, Sheet 13 of 40

FIGURE 2, Sheet 14 of 40

CACCACCATCGAGGCGGGTATAAATAAGGGGCGCTGGCACATGGGTGGCACACTCGCATCATGTCTTCCAACCGATTACGTACGTGAGGGGCAACGAGG  
AGATGGTTGGGGACTATTCAAGATGGACAACCTGTCAAGAACAGGAGGAACAGACAGCAACAGTATTCCCATAGTTTCCGTCGCCCAACCAACCAACAA  
TCAAAAAAGAACATCAACCAATTCCTCTCTGCTCCACCTCTCCATTCCTCATATTAGTTGGGGAGCCCTCGGCAGCTCAATGTATCGACTGGAT  
GACCAGTGCAGAAATTCGGATGAACCTGGCTATTACAATTTCCACTCTTATGATAGAAAGAGGGAAGAGTTCGCTCATTAACCAACACTCCAAAGTGAAG  
GCATGTGGCGCGCACAAGTAGATCTTCCCTCTCTTAATAAGAAAGAGGACGTTGACGAAGCTCCACCTCTCAATCAAAACCAACACATGTACCCCT  
CAACAAGTACAGTTTCCGTGAATATACTCTTCAATCAAGCTTGTGAATTTGGCGAGACCCTTCAACAAGAAAAAGGAGCAAGATCTTCAAGAGGAGAA  
GCTCGCGCCCTTACACCACTCCCAAGAAAGGAACAGAAAGTAGAACTAAAGATGATGTTGTATCGAGGAAGAACTGCACCAAGAACAGAACAG  
AACAGCCCCAGTTCCAGACCCAGATATTCGCAATAACTGCACTACTACTACTACTAGTTGCAACACGTCACGACGATTCTTCTACAGTATTTCT  
CAGAAATGTTATCTGAGTATCGTGTTTTGGTTTCTGGGTGTTTATTCTGCATATTGTCAAAATGTATTAGATCTAAGAAAGGAATAAATAAATGGTAT  
ATGAAATTTAAATCTTTATTGTTTCTTCAATTAATCTTCAATTAATGTTCTCTGAGTGCAGCTTCTGGCAACACTCCCATGTCCCAAGTGCTA  
AAGGTTGTATAGTTTCCATCCCTCTTGGTGGGAATGGTCTCTATCTGATTGCAAAACAAAGGGGAATCTATCTTCAATACCCCTCTGAACAGAAAC  
AGTTTTTCTCTGTATAATAATATCTGCTCATTAACTCTTCAATCAAAATGGTGATGGTGCAAAATTTCTTAGATTAGGTCCCCCTACGCTTCTA  
AAGTAGAGGGTTGAAGATCATCAATGGTCTTGGGCTTAACTCTTACCCCTCCCAATGTTGTTTGTTCCTCAATAGATCTGAAGGTCTCATGAAAT  
CTGCATAGCTTATTTCCACACCTCTTGTACAGCACTCTGAGCAGCTGTAGGTGCGCAAGATTGCAATCTCCATGGTAGGGTGTCAATACAGGAAGTAC  
ATGGACTCTTCTCATGACTCTAGGACTCTTTTACTCGCTCTATAGAAACCGACCATGTCTTTTGTCTCTCTTACATACTTTGATGATGATGATT  
CTAACCGGAGATTCTTCTGAGCTTCTTCTGCTTATTGATTGTAAGAACTTCCGCTATCTTCTCAATCCCAATTAATCTGAATAGATAATCTCTCT  
TCTTTGACTACCGGAGTGTCTCTCTTACATTTCCATATTGCTGTTAACTACTTCCACCACAATCTACCTCCATCACTGGTGGTGGTGGTGGTGGT  
TTCATTTATTGTCGAGAACTCGGCGGATATTGTTGGCGAATCTATGCTGTAAAGCAAGGCGTGTGATACAGGCTCACTATCTTACACCTCTCAACG  
TAGGGGAAAAATGTTCCACGTTAAATGTAGAGGCACTAATCCATGAGAAGCTCGAAAAGTCCACTATCTTCTAGCCCTTGCACCTCTTGACCCG  
TTTTTGTGCTATCTCTTAAATGTGGAATTTTCAAGCTCGGAGGAGGACGCGCGGCAAGCGTTCATAATCTCTCAATTAATATGTATTAGATACAAT  
ATCTGTCATACCAACATCCCTAACAAATCTGCGTCTCTATCGAAGTACCTCTCTTCTGTCAAAACAGCAGCAGGTACCTTTCTTGTCTTGTGGAA  
TAAGGAGTGAAGCGTACTCTACGATGGCAAAATTTTATAGGGTCAACATAACATTGGAAGATATATCCCATATTTCTTCTGTCAGCTTTTAA  
GGTTATCTTCTTCCATAGTTTATGTTTCTCCGATAGTATTAGTATTCTAGCGCTCTCGGATTGAGCTCTCAGGAGTGTCTCCAGAGATGAACATC  
GTTTTTAGGGAGGTAGTCATGAGGGGAAACGTAATGCTTATCTCTGCTAGGTGTTGCGGCTATAGGGTCACTTTTCTCTCCCTCTCCATTC  
ACTTTGGGACAGAAATAACCGTGTGGTCAATTACAGTTCTCTTTCTCCCAATTTCCCGATTATTCGAATATGTTTTACTATTAGGATCTTGATCT  
CAAAAATACTCTTTCTCAATATTGGCAGGAAGTAGCGTGTGTTGTTCTCTACATAACTCTTTATAGAAAATTTAGATAACGAAGATGTGATTGCGG  
GCCAGGGGTTTGGGACCATATTGTAGTGTACTGAGAGAGGGCACCAGCAGCGGTATCCGAAGATTGGCCAGTTTCTTCTGACACAGGCGGG  
GAGTACAGCGCTTCAAGTAGTGGCGCGCGTGTGGGGCTCAGACCCCAATCTGCTTGAACCGATTAGGGATCAATTTCTGGCCTAATAACGTTTCAG  
AGGAACAGCTTCAACCTGAGTAGCGACTCAAGAGAGCCATAGGAGGAGTGAATTTGAGGCTCTCTGACTTCAATTTGTTGCCA  
TAATTCATCCCTTTCATGGAATTTTAGAGGGACTATTCCATGATGGACCGGAGCGTGGCGTTCGAGACTGATTCTTGGACACTTTTGAATGAGAAA  
CTGGAAGGTATGGTAAAGGAATCCCTTAAACTCCCAAAATGATGAGTTGAGCAGCGCAGACACTCCCGAGGAGGAGGAGGCTTAGAAA  
AGGGTCGTGCTGGAGCGTACGAATTGAAGGAGGTTTAAATGCTGTGTTGCTGAATTTAGAAATAGCTCTAATGTCTCTCTTCTGATCATAGTGT  
GTTGATGGCTCTCAAGTTGTGCGCCGAACAGGGTCTCTCTCGCGTTGACCGCGGATCTGCGTTGCAACCGGTGTGTTGCGGAGAGACATGACGGA  
ATTTTCATAGCTCTTCAAGAGGTGAAGATGCACCATTCATTAACCTCCATCTCCAGTGAAGATTGCTCTGCAAAAATCTGGGCTTTTTTGTATAACG  
TTTTTGGACCTTTCCGCTCGGCATACCCGACTCTTCTTATTTAGGTTTATGCTTATTTAGCTCTCTCCGAGAAAATGCTCTTATTAGAGGCTTTC  
CGAGTTTGGAGGGTGAGCAAACTTCATATGCTTCTTTTCCCGCTTCTTAAACACTGAGATGAACAGGCTTGTAAATGATCAATATTTTGGGA  
GAATTTGTAGACAAATTTTCTCACTGAATCTTGAAGGAGGTAACCTTTCTTACACCTCTGCGTGTAAACAGTTTGCAGATAATACCTCAATTTCTT  
TGAATTTGGCATTGTTTGGTTTCAATTTGATATTGATCTTGTATTGATGGGACCGCAGACCTCTCTTTCTCTCGCGTTCAAGTTGGGGAATAATAGTC  
CAGAAGTAAATTTGTAACGTCGTCAAACTATTACTAGAATTTGTAACCTCATAGAATTTAATGACTGTTTCGCAAAAGATTGGTCTCTTTCTTGTGCC  
GTCAAAAGAGAGTGTCTGAAATTTGCTCGAGTGGAGTGTGGGCGTATCGTTAGCGGTAGGTCTATAGTGTACCAAGGGCAATATCTTATTACTGTTT  
TATCATTTTATAGCAAGGTTTTTCTTTTACTCTATGAGTGAAGGAACAGCAGTATGCTGTTGCGCATATGCCAGACTCTGAGATTTTCTCAGTTT  
CTGTAAGGGGTATCTTTGTACGAATCTGACCTGCTTGTACAGATTTTAAGGAAGTATTTGATAGTTCGTATGCTGTTGCGCATCTCTCTGTC  
GTCTTTACTCTGAAGGATCGCTGCTCTCATGTGAACCGTCTCGGGGAAGATTTTACCTGGGAGCTATTATGCCCAAGCATGATCTATGCTATCGA  
TTCTTAAACCGGGAATATTGCTATGGGTAAGGACGACCTTGGAGGTAGCTTGTGATGTTTCAAGTAAATTTCCCTCCAGCTCCAACAGGCGCAT  
ATGCTTTAGTGCATCTCCGTTCCAGATGTGAGCGGCCACACTCTCTGTTTCACTGTAGGGCTACACCTCTTACAGATTCAGTCTGATGATTTTCT  
TGTTCTGAGATCTTCAATTTCTGAGTGCATTTGCGGCGGCTCATGGCAGATTCGCGCAATATCTCTCCCTCAATCAGATGGCTCCGTTAAACG  
CTGTAATTTCAAGTGTCTTCCAAAAGAGAAGCGGCATAGAATAGCCAGAGCCGACATCTCTGATCGTTTCTTGGTCTAGGGAGGACGCGTCTCT  
AGTACTACCACTTAAAGTTTCTGACCTTTCTGTTTATGCAAGTAAATTTTGAACCTCTACAAAGGGTTGAGCAATTTCTTGTGCGCATCGGA  
AGATCTATACCGGCACTTCTCCAGTCTGATTGTGACTTTTCACTCAAAAGACGCGCTTGTTCAGTATCTCTTCAACCCCATCTGATGATCTGAT  
CAGGGACGCTTCAATTTGTTCTTTCTGATACACCTCTCTTCTTCAATTTTGGGCGTGTGTTGCGGAGGCTCTCTGCTGGGATAGCCAGC  
TTCCGACAAACCAATCATAGTAGAAGTCACTGCTTATGTGAAATCGATATCTGCCACATCATCAATATGAGAATGTTGTAGGGTGTATAGGTGCG  
CCCACTCTTTCTCTCACACAGTACAATATCTGTGACCCCAAGAAATGGGCTCTAATCTTGGATCTTATGAGAAGGGAAGGATCTTTCGACACCAAG  
ACGAATCTACACACCTCTTTTTTGGTGAATAAGTCTGTTCCCTTTTCTCCCTCAGTTATGACTCTCCAGTATGACTCTCCAGATTCAGTCTGATGAT  
GACAGTAGTAAGAGGGTTGACGTTTCTGTAGGCTGCACTGTAGCAACATCATGCTGCTGTTGCGAAGCGAGAAGGCTGAAAAATATCTGCGG  
AGAGGAGACAAATCCGTCATATTCTATCATCTCTTCTATACCTTCACTGTTTCAATTTCTTACTGCTATTGACCAAGTTTCCCTCCGAGAGAAC  
GACAGCTAATGCAATTCAAAATCGTTGCCCTCTACGCGACATTTTTTAGGAGTTTTTAAACAGTCAATGTGTTATAATGGGAACATTTTCCCTGGTAA  
TAGAGATTCTACAAAACCTTTCCCTATGATGCTGACTTGAACCTTACCGGAGGAGGCGCTCCGATCGACATATAAGGGGTGCAAGGGGTGGGA  
ATAATTTCCCACTTTTCTTTTAACTTTGACGCGTAAAGAGTTTCAGAGACATGGCTCTGTATCTACAAAAGAGCTCGTTAAACACAGTTGCACTCAC  
CCGGGTACAAATAGCAAAAATTAATAATTTGGGAAGACGATACAGTCTTTCATGACAAAATACCCCGAGAACAAACATTTGTTGTCTAGGAACAAAGAA  
ACATTAATAATGTTGCCGAAGCTCCAAGACCAATCCCGCGTGAAGCAAGAAGAGGTCCACCATGCTGGACGCATCTCCAAGCGGAGGAGCCCAT  
CAATGAAGAAGCGTGCAGGAAGAAGAGCTCCACTGTCCGTCGCCGTTCTCAAGAGCGGAAAGAGTCTGGAGCCCGCAAGTCAAGGCGTTAATTTCT  
CCCTGTACAAACCAATGTTATTTAATGATTTTTTCTTCTGTAATAATTTGGAATAATAAAACATCCATTGAACTTATGAGATTTTTTATCAAT  
TTTTTAACTCACTATAATTTCCACATGTGGTATAAAGTTTAAAGGATACAGATATTAATAGATGATGAACAGTACAGGCTCAATAACACAGGCGCTG  
AAGAAGTGTCTATGAATTTCTGTAGAAATCTTCCCTGTCTGACTTGAACGCTACATTTGGCAATACCAGAGGGTCTACTAGTACACCTGAGCTAGTGCC  
GGGTTCTATTGTAGCATGCAATTTGAGGGATTTTAAACACATAACATATCATCATGTTGCAATGCTGCTGTTTCCAGAGACGCGGATGATGCT  
GAATTTCTTCCCTTAAACTAGACCATGCTCATTTTCAATGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT  
ACGACACCATCTGCTTGTGATTCCCACTATAGCAAAATAGTAACGAGGAAGTAAATTCATTGTTGCACCATATTATTGTAGAACCGGTTTGCATACC  
AATTCCTCCACATCTAGGCTGAGCAGAAATTTGCCATCGGCTACCCAGTACAGGAGTACAGGCTGCGGCTAGGCACTGAGGCTGAGTGA  
GGGGCAACTGGTGTCTCAATTGAGAATAAGAACCCATGCGCGTAGGCGGCGAGTGTCTTCCAGTTCGTGATCTGCTGAGGTTCTGCTGAGGTTAT  
TATTTGCTCAAAAACAAATACCCCATTTTCTTAAATAGCCGAGGTTCTTCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT  
GTCTACCACTATGATTTGAAGTGTGCAAAATTTGGTCTAATGTTCTTCAAGAGGCTCCCTTATCTCTTATAGAAATGGGAGGATGTAAGTGAACAG  
TAGACATGCTCTTGTGGTTATGTCGGTGGCGGGTGTGCTCTTACAGGTGCGCAAGGTTGAGGGATCATTCATAGCAAAATCTGTTAGCTGTTA  
CAGACATGCTGAAACCTGTCTTCAAAATGCTGCAATGCTTCAAGGCGCTTCTGATGCTCACTCCCAATCTTCTTCTTCTTCTTCTTCTTCTTCT  
AACCCTATGCTCTCTGACCGCTCAATAAATATAGGGCGGCTGTTTATCTCGAAGAAAGACGTTTCTTTTGAAGTATAGAACAGACAGCGGATCG  
GAGACTGTGATGTTGATGCTTTATCTTCAAGGACAACTCATCATCATCATTTGACGATAACATCAGATACGGTCAAGACAGTACTTTTGTCTCTG

FIGURE 2, Sheet 16 of 40

FIGURE 2, Sheet 17 of 40







21/201

FIGURE 2, Sheet 20 of 40





FIGURE 2, Sheet 22 of 40

FIGURE 2, Sheet 23 of 40

FIGURE 2, Sheet 24 of 40

FIGURE 2, Sheet 25 of 40



27/201

GATAACATGTGTTTATCGTTTATTTGGTACACTAAATGAAGTTGAAGAGTCGGTAAAGGCAATAGTATTGTACTTGAAAGAAAATGATGTATGATAAAAAAC  
CAGGTATATATCTTCCACCCTCTCTTTTAAATTTTCAGTTATGAATCTTCCCTGGGTTCATGATCCGCTATTTACATCGTATACATTTTTCAT  
AACCACTGGTATTCCTTTCTCAAAAATTTGCTTCACAGATAACAGGATTGAAAGCGTCTTCGTAAGAGATGTTCAAAAGTCATTACATTTAAATGGCTA  
TTGATTTCTTGCTTACTTGTATCTCTATATATTTTATAGAGGCAATTTTGTCTACCGAATAGGCAATGACACCTAAGCGATACCTTTAATTTGGTCACTCT  
TTCTTGTGATCCAAGAATTTGTCATGGTGAAGCAGCTCCCAATTTGAGCCTCCACTGAAGCGATTTACGCTTCAGTCACAGCAGGCTCATCTCCGTTAGACGA  
CATTATTTACTCTCTCTTAAAGCAGTGATTGCAAAACATATATCGTCCCTTGTATCGCCTCTTTTATACGTTTAAACAGCGTATGTCAGATTCTCCG  
TTAGTATATGGGTGTGTTTATATCTAGATTTACATACAATATCAATATCTTTACTCTTCCCTTCAACAATCAGATTAAAGAAATGACTACCGCTTCTC  
CTGTTTCGATTAATGCTTCAAAAGCTTAGCACGCTATTAACTACCTCAGTACGAGAGACAAAGAAATCAGTGCATCTACATCTCTATGGAAGCAAA  
CAATTTTGGCGTTATTTTAAATCTCCCAAGGGAATGGAGAAATCGTACCTTTGTAATTTCTCCATCGTCAACCAATTTCTCAATTTGTACGTAGCAACA  
CTACACCGCCCTACTTGACCAGTTTCTTGTGAATTTTTCGCTATATCTGTGCAAGATATTTGGGGCCGATACCAATGTACAAGATTTTGGCTTAGTT  
TTCGTTTAGCCAAGTAATCAGTATTTATTTTGTCTAACATCTATTCTTGTGTTTCAATTTGCTTTTAAACACAACTCTTTTATGTCTTCTAGCGCAGA  
TCTCAGTCCAGTTCTGCATCTCTATGTAACCTCTGCTCCATGTCTTACAGTGTCTTCAACAATTCGTTTATGTGTTTGGTTATGGGGTGTATCAAC  
AATTCAGATAGATTAGAAACACCACTCTTTCTACAAATTTTCAAGGATGAGGATCAACCTCTCATCGTTATAGTTGGCCCATTAGACATCTGTTGTC  
TACTTGTCTTTGGTATACACATCAACAGGTGTAAAAACATACAAAAGAAATAATAAAATACTTTAGGGGAAGAAAATGTTTAGAAGTATTTTCAACC  
TTTTATGTTCTGGGGGGAAGATTATATACATTCAATCTTTAAGCTAGACCTTCTTCTGCTTCTTCTTACTCTTCTTCTTCTCTCTCTCTCTCTG  
TCCTGGAATCGGAATCGGGCTCGGTAACCTGGAGAAGTTGTGTTGCTGAGTGCAGATATACCTCCTCTATTTTGGTATCTCTGAGTACAAATTTTGG  
GTTTTGGTTTAAATGTACTGTAGAGGGAGTTTCTGTGTTACCTTATAGCTCTGTTTTCAGGAACAGGGTAGGTGTCTATACCAAGGGTAGCCATTTGC  
TCAAGTGTGCTCCACCCCGTGCTTTTCCCTTATATACAATTTCTCAGCTGGTCTTAATAAGGAACCAAAAATATTTCATCAATTTCTTTTATTTGAT  
CAAAGAACATGTTTACACAGTTTGGGCACAGATAATCAGGGTTAAGTCTCAAAACCCATTGGTTCATTACAATGTACTTGACAATATCAGAAATGTCAG  
ATTTTGGAGAAGCATAGGAAGAGCTAAATTTGGTAGAGTGTCTAAGAGCATGCAACGGTTTATTCACAACCATTTGCCAGAAACATGGGAACAGATAGTA  
AGAACTTGTCTCTACCGCAATATCTCTGTATTTGCTCTTGTGGGATGAAGAGGACAAAACACAGGGTATTTCAGATTACTAACCATCAATTTGG  
TACACATACAGGCTTAACATTACAGCAACATAATCAACACAGTTTATCACTCTTTAGTGTTTTAAATGAGGAAGTCTCTCTTACATTACAGTATAC  
GGAGAGCCATTTCAAAACATATGTGTAATGGTGGCAAGAAATACATCAATCCCTAACATCCATTGAACCCAGGCAATCACAAAATGTTTGTAGTGA  
AGAACCACTCTATTAATTTCCCTTCTTATCATATCTTACACATCCATACAGGATTTGGTAACTCATACCAATACAAATTTGCAAAATGATGGCT  
GGCTTATCTTCTGTCAGGAGAGTGCATCTGTATGTTAAATCTCTCCAAAGAAATTTTGGTCTTAAACAGGGCATTTCTACAATAAATCTGAAT  
TGAGCATGCTATTTAGTTGAGCCATGCGCTTGTCTAGAAGCAGCAATATTCATAGCCATTGTATTCAAGATTAGGTTAACTCTGGCAGCATTTGAAG  
AGGTGATGGTAATTTATTGCACATGGACTCGTAATTTGCTGGTTCTTACGGCGCAATTTGTTGCTCTCAGCATCACAACTATTGTCTTCTTCTACTTCA  
GTGATAGTAGCAGAAGAAAGAAAGATTCTCCGTAATTTGTTGAGCCTTAGCACAAGACTCATTGAAGTGTACAAATGGAAAAGGTTAGTATTTCT  
CCTCTAACCGGAAACTGATTTTACGCTCTCAATCTAGGACCTTGTCTGATCAAGTTTTCATTAAACACTGGAATGGAATGATACCGTACT  
AATTCATCTCTCTTTTCTGTCAGTGAATGCCCACTAAGGTCATTCATGCTTTCTTGTGATGATTAATTCGTTGAGCGTGTGACCGTGTTTTATTCATG  
TTTCCAATTTGATCACTTTTCCATCATTAAGATGAATGGATATCTCTCTTATCCAAAGCCGATAATTTTCTTAACATCTCTTCTTAT  
TCTCAAAGACAAATGTTGATACAAAGGAAAAACAATAAGTCTTTTAGACATAGCTTCGTCGACAAAGAAATGTAAGTGTGGAGCATCAATAAT  
AGGAGACACAGTAATGCTGCTTTTGGCGCCTGTTGGAATGCTCTTAATCTAGACGCTGTTTGGGGTGAAGTGTATTTCTTAAATAACTGGAGGGA  
AATTTGGCCCGGGAAGCATGTTTGGTTATAGAACTGGACACATCTTTTGGGGGATGGCCTTAGCGATTTCGGTAAATAGGAAGAAATGCTGAC  
TGTTATGTTTTGGAGTGCAGACCAATTTGCCCTCTTTTAGCGGTATCGACATGAAACAGCCCTTTTATTACTGTCTCTCCACCACTAGAGTTAGA  
AGCAAGGAAAAATGTAAGTTTGTGCTTATTTTCTCAAAAACACTTAAAAAGTTTGAATATCACTGCTGTTATCTGCTGTTTATGTTAGTAAGAGAC  
GATTTTGTGATGAATTAGGAAGAGTCTGCTTTTCTGTCAGAAAGCAAAATTTTATACCATACATGCGAATTAAGGATGCAAGATTTGACCGCATTA  
TCTTATCTCTCTCAATACGCTTCTTAAAGCAAGAAAGGACATTTTCGAGTCAAGAGTGAAGTGTACAGAACCACTGAACCATCAGAAATGTAATTCAGG  
AAACTATATTTTGCAGGTTTCTTAACCATCGAATTAAGATGCGTCTATACCTGTCATGATGTTTCAATAAGTCCAGATCTCTCACTAATGTTT  
GCAGCTTCTGGGCTCTTATTATGGACGATAACTTGGCAGAAATAGTCTTGATGTTACGAATACTGTCAATATTAATATTGACAGAGTAGGGCCAGAA  
GAGCCTTCTCGGCAGCATCATCGACAGCTACAAGTGTGTTTCTTGGGAGGGGAAGTAAACCTGGGTAGTTGTTTGTGAAAGATCATCAGAAATTTT  
TCCAGTGTAAACGAATGGTACGCAATCAATACACAGGGGGCGATGTATTACCATCAATAACTGACTTTTAAAGCCCGCTTAGTCGACAGGCAATTA  
TCCCCACAGGATTTTAAATTTGAGCATCTCTCGGCGAGACAAGATGACGTATCATCTCGGAAACCTTTTCAAGTGGCAGGGTTGACAGATTAACAG  
GCAACCTCGCGATCAAAATGATGATTTCCATTTCAAAAGGCCATTTGTTTCTGATGTTTCAATATCGCATCAACTTTGCTCTCAATTTTCATATT  
CTGCTTATTTTCTTCTGTTCTGCTCATAAATTTCTCATTTCCATTCATCTATATCTTCTTACAATCGTTTCCAGGAATGAAGATTATTAACCTGGCAGTT  
TCATCAATAATGGAATCAAAATGATTGTTGCTGCTATATTTGGATCGCAATTTACCTAATATGCATTTCAAGGCAAGTTTACTTCGGATTTAGAGAAAT  
TGCCTTCATATTTCGATAGGTTTCTTACAGTGTCTGATACAGCATCTGCTTTCGTTAAGCCAGGAGAATCTGTCGAGTCTTAAATAGAGAAAGGATGGC  
GCTCTATGAGGATATTATACAAATATTTTCTCTTAAAGGTAGTTGAAACAAACACAGAGAAACAAATTCACCAAGTGAAGCTTTCTGCGCAAGAA  
GAAATATTGATAGGATCTTTATGCGGACACCATTTTCGTCATTTTACCCTGCTTCTCTCAATTTCTTCACTATCTACCCCTTCACTCTCAATTTTCGC  
CACCTCTTTTACGTTTGTGATCATTACAAAGCATTTTCTTCTTCTGTTGAGTTCAGAAATAAATCTGATCTTGGTGGCATATTTTAG  
TTTTGATTCGCTCATTAACCGTTTAAATTAATAAAGCATCTTAGAGTTTCTCCGCTCACCCTAGCGCTTCTGCTCATTGATGTCAACATATTC  
ATTTCTCCCGCACTGCCGCAAAATCGACATTTACCCTTCTGGAAGCAGACTGCTCTTATCGCTTTCGCTCGATGATTTTGAAGCAATGAACCT  
GGTGGCCAACTTCGTCATAATACGATACCTTTTGGCAGCAAGGAAGCCAAATTTGCTTGAGTAGATGAAGAAATCCCGAGCATTATTGAGAGC  
CGTAACGTTAATATTGTAAACCTTTTCAAGCAAGTCTCTTACACAAATCGCAAGAGTGGTCTTAAATTTGTTTCCCTTTGTTCCCAAAATACAAAC  
GCAGTTTGTGGCAATATTCAAGGCAACGGAACAAAGAGTTGAACCATGTGGTGGCTTTCTCTCATCTCGAAGCAATGGCTGATACATCTCTC  
CAAGATCTATTAGGCCAGCTAAAGATACACGACATCCATTCGGGCATGCGAGTTTGGTCAACACATGAGATGAAGATTTCTTAGCGCATGAAG  
GGCTGATTTGTAGAATTAACATCATCTTTTGTCTCTCTTTTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT  
GAGGAGGAGGAGGAAGAGCGTTTCAATCTGCTAGTGGCTTTAGCAAGTTCGTATAAAATGATAAAATATATAATCTTCACTAGCAATAAAAATG  
ATGGAGGTTTATGAGAAGAACTGAATGCTGTTTGTATACAGGTTTCGATTTTACAACGCTAGTATTTTCCGTTGTTTCTGCTCTCATCTCTCTCT  
TTCATTTTCTGTCATCATCTCTCATTTTATCGGTTTCTTCTGATCTTCAGCCATAGACCAAAAGTTTGTGTCATATTACTGACAGAGTCGGTATT  
TTCATATAAGATACGTGTTGATTAACAAAGTCTCTACCCATTGTAGTGGCTTCCATGGACGAAGCTAGAGCAGCAGAAATAAATTCGCTTGAAGATA  
AAGAAATCAAAATACCGTGTTTTCTGCTTTTAAATTTACACCTGTGATACATGTCGCTGACCAAAATTTATCCCTTCAAGCGCTCTCTTAGAAGATAAGT  
ACCATCAATCAAAAAGTACCTGTATTAATCAACGCTAATACAGCGAGTGTGTTTACGTAACCTGAAGGGAATGCTGTACCAATCAACAGTCTTA  
GCATTTCCCATTTTTCATCTCAGCCATAATTTTGAAGATGAGAGCAGTCTCTGCTGATTTGTAATACAAAGACGATATTTTGGCTCTTTTGGCATAT  
TCTTTGGCAGGACCTGGATTAGCTAATAATGACATTTGTATGCTTTGTCGCAAAAATATTCTGCTTGTGATGTCGACAAAATTCGGGACAGTA  
TCTATCGTGAACAACTCGCCTTGTCTGACTCATTGTCAAACTTGAAGGAGTTAATCGCTCGCTGAACATCCCGGTTTCTAAATTTGGAGAGTAGAC  
TCGATCGCCTCTTGTATTCTTACCTTTGATGATGAACAAATACAGGCCCAATCAGTAAAGAAATGCGGGGCAAAATACCTCTCTGTCCTTTGTTCA  
ATAAGAAATCGCAACACTCTTAAAGATTTTGTCTCTAGGAGCGCAAAATTTACCTGACAGGGGAGTTAAAGAGGAATTCACCAAGTAAAGACCC  
CTTATCTGTTCTGCTGTAAGAAAGTAAATATCTTCTTAATGGCCAGACTAGATTCTGCTCAGACTTTTCCATTAATGACAGACCGCATAGAAATTAAC  
ATCTCTTAGTACATCTGAATCTTCTTGCAGAGGCAAGTTTATCTGTAACCTTGTATCGTCTCATCATCTCTTCTGCTGCGCAGCAGCATCAAC  
TGCCAGCATATAGCCAGAAATGGTCTTATGTAATGTGACACGAAGCAGCGCTTATGGCCATCAGAGTTGGTACTTAAACACCGGAGTTTAAAG  
GATACGATCAATTTCAATTTCTGATGAACAAAGTTGGGACATTTCTTAGTTCTTCTTACCAATGAACAGTAAATCATTTGCTTGTGCGCATTAACAA  
GATTTGTTGAGGAAATCTCTCTCTGCTTATATTTGTCTGCTGTTGTAACCAACACTATAGTATGATCCAGAGTAAAGTAAATACCTCGG  
GATCTAAACATGAAGACAAACTTACCGGGCTTAAACAGCCTTAGCCAACTAGCTGCAAGAAATTTTTCGGGAGAGGATCAGAGCAATATTGTCCTT  
AGATTTAGCAGAGAAAGTGGGTGACAGTTCATAAGAGGGAAGTGTCTCTCAGTAGTTGGCCATAGTCTTCCACACTCAGGAACAGAAATGCTATCT  
GAAAGTGGAGGTTGATTCGGTAAACGCCCACTAGGAGCATTTCTTGTGATTTCAAGTGGCTATTCTTGAACCTTGGGATGAATACATCCCACT  
TACACAAACGCTAATGTGACGCTCGTACAAATTAATCCACAAATGTTTAGATTTTAAACAGCAATTTTAAAGCAAACTATCTGCTTCTTAGGGGA  
CATACCATTTGTACAGTTCCACATCTTACCCAACTTTTATAGGTGTTGATATTGGATATTTTGGTGGGTAAGTATTGCTGGGAGAAAGTCTGCTCA













FIGURE 2, Sheet 32 of 40

FIGURE 2. Sheet 33 of 40





CGCTACCGTTGCAGCGGAGGAGGATATCTTCATCTGCTTTATTGGTACCTACATAGGGTCTTGGTGCAGTCAGAAATTTTTTCAGAAGATATCGAT  
TTACCAGAAAACAAAAGAGGCTGATAAGGAAAAGGCATCAAAACCGATACAAATACAAACAAAGCGCTCAACGCCGCGCAGACAAAGGAAGC  
AGAGCCCGCGCCAGAAACAGGACTTTAAGAAAGGAAGAGGAAATATCCAAAAGAAAATACACGCTCAAGAGAAAAGACTTAATTAATCGGCGGAC  
TCTCAGAGGAGGAGGAGGATTAAAGCGAGCATCTAGAGGCAAGGACTGCTTAGAGGGGATTGAAAAAGGACATTTGGATTGTTTTCTCCCTTTTAG  
AAAGGATGTTTTCTTTAATATCATGTTGATTATATTAGCCCCATTAATATTGTTGGGGTGAATTTCTTCGTATACATAAAGAAAGATACTGGTTCTC  
TGGGGCCATTAGTAATGACACTACTATTAACTCCACAAATCTCGAGGAGTAGGAGGATTATGGGATAATGTTCTGATTAGAGTCAGGAGCGTTTCCA  
TCTAGGTACAAATATCCACATCAACGAAAATTTAGTCAATATAGCCATGGATGGAGATTATGCGACATTAGTAAGGAACGGAATGTCACAAATCAAA  
GGCCATATGAAACCTACAAAGACGTCGAGGATAGCCAGTTCTATTACATTTTTCCACGTTAGAAATTTAAACCTTTAAACGGTGATGAAATAAAGA  
CCACTTGAAAGAGGATGAAAGTTTGTGTTGATCGAATCACCATAATTAATGGAGGGTTTATCATACAATATCAACAAACCAATCCCATTTACAAAT  
TCTACTGAAAAGCCGATATTAACACGGAGATAACTTCCATCGTCAGACCACTGGTACAGATGAAAGGTTCTTCTGCTCGAGAAAGCAATCGTCGAAG  
ATGGTGAAGAAGGATTACAGAAAATAGGTACTTTTACGCCACATGGCAAGTAATATTGTTGTAAGGCTAGTTTCAAAATCAGTCATGCCCACTATTGA  
AATTAGTGATCTGACTGCGACATACAAATAAAGAAATAGTGTAAATTAAGCCTGTAACCTGGTACTTCTATAATTTATGAAAATTAACAGGAACAAAA  
TAAATACATTTTAAAGTATTATGTTTATTGTTTATTCTCTTAAATTTTATTTTAGGTATCCTCTTGAAATATGGAACGACGGGTGGGGGTTTTTC  
TTACCTCTCTTTGGTTTTGGTTAGTTTTCTTGGCCCCACCATCTGGTGTTTTTGGTCCCTTTCCCTTTCCCTTTCCAAAAGCCTTTCTGAAATGCAC  
CGCCCACTCCTGGTATTGATCCAGCGCAGAAAGCGATCGTGCCCGCAGTCGTGTGCGCCGAGGAGAGATGCTAAGAGAGGAGCTTGATTTATTATTAC  
AGCAATAACATGCAAAATATTGTTTATTAAGTAAACCAAAAGTATTTCTTCTCTCTCTTAAAGACGCCCAATGGCGAGGACTCTTGCTATCT  
TTATTGTTATTATTGTTATTGTTATTATCCCTTCCATGTTGTTGCACATAAATAGTGTGATTATAACATTAAGAAACATATAAATTTTGGCTC  
AAGGCACCTTCAAGGCTGTGAAATATTACATGCAGATAATGTTGATTCTGAATCAGGGTTGTGATTTTTGGTGCACGTTTACTGACGTCACCAATAG  
GCATATGGATAGTACAAATTTGTTGAATTCCTCATCTCTTGTCTCTTACCATCTGGGTTTGTCTGGGAACAAATGCTCATATTTCTCCCTTTT  
TATATAGTAATTTATCGGGTTGAATACAGTAGTTATGGAATGGAGATGCGAACCACTAAAGCTTGATATTTATCAAAAATTAAGTCACTTTCT  
TCGTATCCACTCTTCTCATTTTACGGTAAATATCTCTCCGTTAATGACGTAGATAATGTTGCATTCAAGACCATGATGTGCTTTTCATGATATTCGT  
TTAAACCAATTTTAAATGATTTTTAAAGGGAATACTATAAGAGACCTCTCTCTAATACGATATTACTGATGGTCTGCTACTACTTAAAGTCAT  
GATTGTGAACCTCTCAAAAGAGCTATCCATTTTTTTTACCTGTAAGATATCGTCAACCTTCCGCTTCCGCTTACCCCTCTCCACTCCCTCTCTCT  
CTCCTCTCCCGCAGTCTCTGCTTTTATTATCGCTTCCACCTAAGATCTTCTTCCACCATTTATCGAAAATACAAAAAACCTACTGCAACAGAGC  
AGCTATTATCAGAAAATCACCAGGATCAAAACAGGGGCTTCGTATTAGACATCTTGTCTACATGGTATTGGAATTTTCCACCTTTAATACATAAA  
AATTTGGTTACAAATAGTTAAAAATATGTAACCATGCAGTACATAAAGGTGATCACTGGTTTTTCATTTCTATATCGTAATGAAAATTTTGCATTAC  
CGGTATCTTTGTAACATATCGGAATATCTTACATCCATACAGCAGATGAACCTCTTTTTCAGGCTTAATTTGCGATATCTTTTCAGATTCATTT  
GTGGAAGATCTTTTACAGGAAGATTTATCCGTTGATTATTGAACAAAAATCGAGCGGTTTTTCCCATTTATCTACTCAAAATCTCAATTTGTGTA  
GAAAATCCCTCTTCCAGACTACAAATCTTGGGCTTAGCCAGTGTCTGATTGTTGCTGAGCGGATTATTTTCTCTCCATGAGGAGATAAATGTAGC  
TGGTTTCATACCTCCCTTGGTCATCAGATTTTACTCTAAAAATGGCCACTGTAATACTACTCATGTGAAGATTTTCAATCAATTGAGGAATCTCTCTCC  
TATATCAATTTCAACGCTAGCTGGCGTTTTTAGTACGTCCTCATGCAACCGATTTAGGGCTCTTGTTCGGTACAAAAAAGTTAATGCCGATTTCCGGATTA  
GGGAAAAACATCCGACCACTCAGTTACCCAGGCACTAATGTTGTCTTCAGACATTTGTAAGTTGCTGCTGCTTTTCTATTATTAGTGTCTATTA  
TCCCTCTGTTGCCCTGCAATTTATTATTGAATACAGTCTCTCCGTAGGGTTAATATGCTCTCCATACAAATCTACCCAGCCATGGAGCAGTTGCACT  
ATCTGACGACCTCGTCCAAACATTTGTCACCTAGAGGAAATTTCCCATCTTTAAATAGTTCTGCAAGTCTGCGCTCCATTTATTTATTTAAT  
GTGTACAAATCTTCAATTAATGATTAAATTTGAAACAAAAGGTAAATAACAAATTAATTAAGAAAGATTCAAAAAGTTTATTCATTTATCTTACACACA  
TATTTTGTGAGGCAAGTACATTAACATTTATCAACATTTCAACAGACCACTACCATCTATTATTAACAAACAGAAAATGGGGAAGTTTGTGACAAAG  
AATCTCCATACATCTCTTCTTACATGAATCAAGAAAGTATTTCCATCATCCCATTTGCTCATCGAGCGCTGTACCCCAACCACTGAGTTCAATATAGAC  
TATTAATTTATCCCATGAACAGTATTTCTGGAGAGAAATGAAGGGTCTTTCCGCTCCAGGTCATTTGTTCCAACTTTTCCACTACTTTTACTTCCCC  
GAGTCCAATTAGATTGACTGTCGTCTTTCTTTTGTCTCGATCGCACACTCGCCAAATGGCAGCAACAAAGGAATCGGGAAGATCACTCTCATTTCTTT  
GGCGGTGAAGAGGAGGAGTACCTGAAGAGGCTGCTTCTATTATTGTTGTCAGTTTATTTCTCGGCAATTCGATTTCTTGGATTGTTGATTGTA  
AGAAATAGTCGATATGTAAGTAACTCTTACCCCTCTCCAGTTCTTCTATACATCACTACCCAGATGTCCTGATGGCTTCTTCCACCATGCGCAAGGAAG  
TGCTCCACATCTCAATCAAAATCCATCAGCAGTTTCTTCCCTCTCTTCCCTTCTGCGAGCAGTTTGTGGCGTACCTGCAACCGATTTCTGCTGTC  
GGCTTGTCCGACCAAGAGTAGACATCATGTACTTGGATTGTCAGCAATCAGATTTACACATACAGCACTGGAATATCAATATTGCCCTCATTCATA  
GCTTGTAGAGAGCAAGTCTTCCATACAAAATCCCAACAGCCTTCCGTCGAGCTGAAACAGCAATTTCAATGCTTTTATCTGTTGGGGGAAGATTGAG  
AGCTCCAAATGATATCACTAGCCATCTCTTACTTTGTCTCAATTTTCTTGGCATTCAAAAGATTGAGGGTTGTGGGTGATGCAATTTGGCAGGAGGA  
TCTCTGGTCAATTTATGTCAGTATAGACCAATTGTTTGGCCATCTCTGATTTTCCCTTGCATCTCAAGCACAATTTTTCATCATTATCATATTTCCA  
TCTTTAGAGATTTTACACCACTCTTTTGCATATATATTCAAGTCTTCTCTGATAGGTCAAGTCGAGCTACAGGGAACGCTCATATTTCTCCCAAT  
CTTTGAGTCGCCAGCTTTTGAAGATTCTCAGAGCAAAAGGCTTCTAGTTTATTTGGTAAAGTTCTTGTCACTCTATTTATTTGGCCATGCTAAT  
AAGCACTCAAAAATACTACACCGGAAGATTTATCTCGATCTCTCCATCGGCAGTTCTTCAAGAAATAGAGGCAAAATATATCAAAATCACTCTCGG  
AATGCTATCTGCTCTGTTGTAATCTTTGTGAATGCCATTTCTCATAGCAGGATCATGTTGAAAGTTGTGGAGGTACTTACAGAGGAGGCCCA  
AGAGCGGCTGCAACACAGATTTGATTGCTGCAATACCTCTTCTGCGAGCTCAGACGTACATCATACAAAGTTCCCAATTTTGAATCAGATCTCC  
AAGTCGATAAATCATCTTCTGCCACATCTTCTCGTCTCCAAATAGATCGCTGGCCAATTTCTTCTGGAAGCACTTCTTCTGTCATCTGATCCCAAA  
AGACTGGGATGGCAAAATATCTGTTGGAGTTTGTGACGTCTCCAAAGCACTTCAACATTTCTAGTAATCTTTAGACATGATACCAATTTCTGAT  
GAAATCAAGCACTGTCAACTCGAAGCTCTGTTATCTGTCAGTGAATAATATCTCTCTCTGATATGAATTTGGAGAACTAAATGACTGGTGGAA  
ATTTTGGCAGCTCTTCCAAAGGATCATTTGGGCATAGAACCATAAACATGAGGATTGTTCAAGTGTGAGATTGTTATGTTAATCAATTTGGAGTG  
CATCATAGCTAGCAGATTTATGTCCTTCTCTTTTAGGAAGGCATATTCTGCTTGTTCGCCATTGAGTTCCAGATCAGAAAGGATGGTAGAGAC  
CTCGAGCAGCTCTGCAATTTGCTGTGAACAAAGGCATAAATAGCCCCAATTTATCAGGCACAGCACTGCAAGCGTACATGATTGTTGCTTGT  
ATCATCATTTTGAAGAGTTTTCAGCAAAACATCTTGGCCAGTGTGTCACGCGACATCTTGCAGACAACTTGTGCTTCTCTGTTGATGATGATG  
CAACACTTGGCATCAATTTGTACAACCTTACAATCAGAGGTTTGTATGGTCAATTTCAAAATCGTCCAACGACTGGAACACGCGCTCATGTGCA  
TGCCAGAGTAGAGGCGGCTTTCTTCCCTGTAATTTGACTGGGCCCTTTGAGTCACTACAGCCAGTTTGTGATACCAACCAACCGCTAGAAGACTGT  
TGTGTTGTTGTTGCTGCTGTGGATTAAAGGTTTGTAGAGCAACAAAGGCATATCATATGCAAAATCAAGAACACGCTACGCTTGTGTTGTTAT  
ATCTTCTCTCTCTCTGTAACAATCTTCCAAAATCAAAAGGTGCAATTTTAAACCTTAGACTTGAAGATTACAGCTTGAGCGCTCGCCATATCTTTCAAA  
TCCCTGAGGCTGATGTTCTTCTTCAATTTTCAATTAAGGATTTCTTCTTGTGATTTCCTTCAAGATGGGGAAGGCACTAAATGATCATTCAGG  
TTCTTTGAAGAGGGAATAGATTGCTCATATAAAGGTTGCGATATTCTTGGCACATTGCACGATTTCTGTCAAAGGATTCAAAAGTTGCCATCTGT  
TGTGTCGTATCCAGAAAATGATCAGATCTTCTTGGCAAAATCACTCTGCTTACTGCACTACATCTAGAGAAATGAGGATGCGTAGCCAACTATC  
ATCTCCATAGGCGCTTCTCATGTCAAGAAATAGTCAAGGTTTTTGTCCCAAGTCTTCTCATGTAATTTCTTCTGTTGCGATGTTTCCGAATGCTAG  
GAGTTGGGACATTAATTTCCAGTGTCAATTTGGCAATACGCACTCTCATGCGTGAACCAAAATAGGAGTGGGAAATTTTCACTCTCTGTAATGC  
CATGCAATTTCTTGAATGAGGATCATGTTGTTAAACACGTACATCAGGGAAGATGAAGGCTTCACTGCCACCGCCGCTCCGCACTCTGCTGT  
CTTGGCATATGCTTCAAAAATCTCCCTCTTCCCACTCTGAACAGTCTGGTGCACGCAACAGGTCCTTTCAATGAGTGTGTTGTCGAATCATTTG  
TAGTAACTATAGCATTTTGAATGCTGGAGCAATCAGATGAGTGTAGCCATGATGCGCAGAACATATCAGCCACAGCGGATTTCACTATCTCCATTT  
AACCATAATTTCCATCAACGAGATCTTAGACAGTAGGTTTTCGATGTTGAGATCTGTATCAGCAGCAGCAGCTTTTATGCTCATGTTCTTAA  
AAGTTCAATCTTGGAGGAAGGTGCATTTTCCATCAAGAGATTAAAAAAGATGAGATTTCAAGTAGAATAGAACTTCTGCTCATGTTGGAAGAGG  
CTTGAACGGATACGCGCCCAACATCTGTAGTATATCTTTATGTTAGAAATGATGTTGACGGTAGGAACAGATCTCGCAACTGTATGCTGAGGTGCTCG  
CAGGCGGCTTTATATCCCTCTTCTGTTGACGTCAATTTGACGCTGTACGATTAAGAGGCTTATTTTATAGAGATGACATTTTCAAAAACGTTTATATA  
ATGTTGCAATTTCTTATTAACATAGTCAATCTCGGCTAAGACAGGATGTTTTTGGCCAAATACGTTTCTGCGCAATCTGACATTTGGCCGACCC  
AGCGGTCACCCCTTCCAACTTGCATCAGGCGACCCAGCGGCTCACCCCTTAAGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT  
GAGGTTGAAAAGTGGGTACCTACTATACACTCAGAGGTTGGCAGATCGGCCCTGTTCCAGAAATGGCTGTCCAGAAATCGGGTCGGACAGATTCCAGAA  
ACGTTTCTAACCATTCTTGGAAACAGTCATTTCTGGAAGGGTTACAATTTCTTATAACTGGTGCATTTATTCTAGCATATTTCTAGCCCTTTGTGCAAT

FIGURE 2, Sheet 36 of 40



FIGURE 2, Sheet 37 of 40

GAGTATGATGATGGTCCCGGAAATCATATGTGGCCAGCAAAATTTACCTGTAACCTTTATTGCGGGTTCCAGTGGAGAAACACCCACCATTAGTACGTGCA  
GAGATGCTGTAACCTTTCTAGGAAGAGCAGCCAGGAAAGAAATGGCAGGATGGGATGATCAATCGGCAGTGGGGCAAGCCATTATAGCGCTAGCCAACTG  
GAGAAAGATGGGGAATTCGCCAAAATATGTTTATTACTAGAAAGGGTAAATGCCGACTTTATAGAGGCGACAGTTCTTGTGTACGTCCGATAAAC  
TACCCCTGTGTATTGGTAGATCCATGAGCCCTAGCTGGAACCTCGTAAAAGAAAGGTGAATAAAATGCTTTAATAAAGGCCCTCTTCCACGAGAAAA  
GGGTGCGTCCAGACGCATCTAAAGTTACTTGAATGGGCAAGAACTATTAGTCAAAAGTTATCTCATGGAAGTTTACTTCCAGCCGAGAAATGGTGCAT  
ACACCGGCCCTTCACTTTGTAGGCAAACTCTCCTCATTACTGACGAAATGGGTTACATCGCTCCAGATGATGCCACAAGAAAGCCCTATATCCAGAAC  
CTAAATGCGGCCAGACGAATGCGGCTGCTGACGCTCATTCTTGGTTCGCTCCCAAACCTGAATTTGCTCCCTGCAAGAAAGGACGATTGAATGGA  
TGATGAAAAGGCAATGATGATGTTAGAGTTGTAAATTTGCTTCTGTGAAAAGGCTATCCAGAAATATGGAGGTTGTGTGAATGTTTGTGAATG  
TGGAACAAACATGTGCTGGATATGTGAAGAGAAGGTTTCTCCTGCTGATTCTAATCATTGTGTGAGAAACACAGGATTGTTTATAGTAACGTGTTAGG  
GTTAAATATGCTTTAGAAAGTATGTACGGGTTTGAGATTGTACCATGAAATATGAGAAGAGGAGTTAAATATTTATGTAATGGAGAAATGGATTGTT  
TCTTTGATGTACAGAAATGGTTGCTAAGAAATATAAAACACTGTGAAATGTTATAACTGTTTATTGTCCATTATATTCAACAACTTTAATTTCT  
TTGCATGCTTGTTCAAAAGGGCAACATAATCGCCTCTTCTCCTCCTCTTCTCTTGTCTCCACTGCTTGTGGCTACTCTTCTCTCTCCAC  
GTTCTCGTTGACGTCTCAATTGCTTAACTGTGTGAAAATCATGCGACGCTCTCGGCGCAGACGCTTCTTCTCTCTCTGGAATGTTTCTTCTCT  
TTTTAATATGCTTGCCTCTTTAGAAAGAGGGTCAATCACTTTTATTCTACACCATCTTGATATACTTTCCATTATCCATGAAGATTGTTGATGCT  
GCTCCAGTAACCTCTCTGACCATCTCCATCAACAACTTTTCACTCGGAGGATCGATGAATGATATTTCAATACCTTTCCGGATGCAATGAATGGTTAT  
TAAATATGTTTCTGATATCTCTGATGCAATAGGTGCTCTGAGGACGATTGAGACATGACCAATCTGGGCGCTACGTTCTATCCAGCTCCCTTTCT  
GTAGTCTGCGAAGAGACGATCTGATTCTTCTCTGTGGCAGAGTAAAAGAACCATTTCCATTGGTCTGCACAACAAGGAGAAATCTCGGTCAGAA  
AAGGCAAGAGACGAGACGCGCCGCTCATATCTGGATGATGAATGGAGGACTTTTATGACATCTCGAAGAGAAATCTTGTCTCTCTGAGGACG  
CATATAGCTTTTGGATTCTTTAAGTTTCTTCCAGAACACTTCTTCCGATCTAGAAAGTTGACTTCACTTTTGGCAAACTGATTCTCTTTTCGAA  
ATTAATGATTTTGGATTACTCTTTCTAGCGGTATACAACTCGAGGATGAAAAGGACATGACATGTAATTTCCAATGCTTTTACGTGACGGGTGTGTG  
CTGTGTGAAAAGGTAATCCAAGTTTATCCATTGGTACAAATAGGACATATTTGAAAACCTGCCAACTATGGGTTTGGCATTTAAGCCATTGTACG  
TTAACCTATCACAGATATCATACACTGATACAGAGAAGAGGTACCACCGTGTGCTCACTAACACTTGTTCATATTCTTAAGAGTACGGGTACAATTTCT  
AGGAGAAAACCCACACCTCTAAGAAAGTCTGTGTGATTATCATATAATCACTTTCTTAACCGCCGCTTGAAGAGCGTTCATGTATGATTCTGTAT  
AGGGGTGAAAATTTAAAGTTCTTGTCTAATACTAGAGCCTTTCTTCCCAATTTTGTTCATTTCTGCAATGGGCTGGTAGTGATTCTGCTCCATTTT  
TTCCATAATTAGGGTCTTCTATCTTCTAAGAACTACCTTGTTTGGGAGACTGGATGTGATTTTATGCTCAATCTGAAGCGTAGGATCTCTGTTAATTG  
TCTCAATCTTCTGTTACCGTCTTGTATGTGGAGTACTTCTTGTCTCTAAGACGAACTATAAAATGCTTCTTCTTCTTCTGACCTCTTCTG  
ACTGATTGTGAATGTTAATAGTACCCGTTCTAATTTTCTGATTTTCTCCAGCAACAACAAATACCATTTTTACTTTTAAATGTTTAAAGCTTCTCCAT  
TTATACCGTGTAACTTAAATGTTGTGTGCTGATATAATAATCTACGAGCACTGCTCTGTATCTGGGAGTGTACTCTGATGTAATTTGCTA  
TTTCTTATTAATTTTCCATTACAGGAAAAGATGATAAAATGAAGAGCTGGGATGAGATATGTGTTTTTCAATATCTGATATCAAAATTTGGGCAAAAT  
ACATATTTTATAGCTCAACACTCATCTTCTGCAAGACTTTTGGAGTGAATTAATAAAAGGCTATATTTATGTAACAATAAAACACATGTGCA  
GACATTAATTTGGTTTGTCTACCCCGCCCAACATCGTTTCTCACTTCTTCTCACTCTATATAAAACAGCTGACCCCTCGGAACCCACAGCTG  
ACTGTGATTCCTCGCTCGAGATGACATCTCCAGCTCCAGCTCCCTTCCACCCCAATCTGATACCACCTTGAACCCGATGTGTTTCTCTCTCT  
GACAACTACAGGAAGTGGTCTATCTACGACCAATTTCAAAATCAAGTGTGAAACCCCAAAATCTGGAACATTAATGGTGTACTTCTGAGTCTGTGAT  
ATGTTGTTTACCCAGATATCCCGCAGCAGATATTTGGGAACTATATGTTCAATTAATCTGCTCAATCTGCTCAATCTGCTCAATCTGCTCAATCTG  
TAACAACCTGTCTGGACAATATTACAGCCCTTCCAGAGTGATTGAAAGTGTGATAAAGAAACGTCAGCGGCTTTAAATGAGTACACAGCTTCTG  
ATGACCAACACCGGAAATATCTATTTCTGTTGGAAGCAAAATTTGAGCAAAATTTAGTAGTGGTGTACAAACCCCTATACGACAGTGGGTATGCACTCCG  
TCTAGAAAACCTTTGTCAAGGAAGAACATGAACTGTCTGGTGACACCACTTCTGGAATGACTGGATTCAACATTTAATAGTTCCTCCCGTGTATTT  
GAAGTGCACATGAGATGGACGCCCTAATTTTATGGCGGCTTCTTGAAGCACAATAGTTTATGGGAGAAATTAACGCCAATATGGAATGTGTACAGCT  
TTGATTATGGCGGCTTCTTCTGGACGAAGATGGTGGCCACCAGGAGAGAGTTTCTGCTGCTCGGAGCACAACCTATCAACTCGTATTTACAAGTGCAG  
GAGAAATATCATGCAAGCTTGGACAATACTACACCAACAAGAAATAGAGAGGAGGAAGAAATGTTGGTGGAGCCTCGCTTCACTTTATGAGCGGGGAC  
GGAGAGGAGGAGAAAGGAAGCCCTAGAGCTAGTTTCTGATGTGATTGGGAGCAACAGAGGAGGAAGATTTGGTGTGATTCAACCAATGCCCCATTTCT  
CAGCCATCACTAAGCTTGGACAATGAAGGAACATGGAATGATTGCTGCTTCCCATCAATGTTCTTTGTATGGAGAACCCAGGTGATGAATCTTCT  
CTTCATATCAACGGATGCTCTAAAATTTGGACAAGCGCAAGCATGGATAGATGAACGACTACGAACCAATGAAATGGAGGAGAAAGAAATTAATGTCTTT  
AAAAAGACCTTCCATATGCTGGCTGATATTACCCAAAAGGCTCATGAACCTGCTTATCCAATACCATCCCATTTGGACCAATGGCAGGAGGTGAATTT  
GGCTGATCTACACTGTGGAACCTATTGCCCATGAATTTGTTACCATTCTCTAGTAAACACATTTGAAATCTAGGGGATTAAGCAATCTCCCGATTCTCA  
TTTTGATACCTTTGTCAACCTTCTTAAATCCATTGGAAGAAATGTTGCTAGTGTATTGCAAAATTTGACATTTTACTGGACATAAACCAATGAAAT  
GTGTTGCTCGAGGTTCTGCTTCTGGGAAGTGGTGGACTATTAAATTTTGTGGGTGTGAACATGCTGGACTTTTCAAGTAAACAAATGTAAAGTGAAGG  
ATAGAAAATATCCGATTTTGGCCTGTATGGAACCTCTCCCTGCTCACTAATCCAGGAAGCAATACCGTCTGATGACAGAAATAGTTTAAAGGGATTCTG  
TAGAGGGGAAATCTAGGAGGTGTAGGTGAAGTGTATCCGACATTCACAGAGGTGTCAAGAAATTTTGTCTCATGTTGAAATAGGAAATTTAGTGTG  
GATAAGAAACCTTGGTTTCACTCTCTTCAGAACTGATAGTGTCTGATCTCTTTTCACTAGAAAGTGAAGTGTGCTGATGATTTTAAAGTGTGATGCA  
TTAATAATGGCCAGTGTAGTGTCTGTATGAGGATCCTAAGTCAAGTGAAGGTGCTGCTGTATGGTTGGCCAAAGGATGAAATGCCATCATCTTTGA  
AAACCTTAACCAAGATACGGCCATCTACGAGCGCTATGGAGCGAGTATAGGCGAGCACAAGATAGTGTACTATGATATTGAACCAACAGATAAAGAT  
TTCACCGACAATAATCAGTCTACATCTATTGGGTTCTGTTTGTGACGGAGCGCATATGACACATGGAGGAGAGATGAGAAATTTTGGAGCTGTTG  
CACCTGGATCCGACGTGGAAGAGGTGAAAGACATATAATAAATTCGTACCATCTGAAGAAAGGAAGACATTATGAAACAGTGCCCTCAAGTGAATGA  
AATTTTCCAAACGAGTTTGAATGTTGCTTGGTTTGGAAAGATATAGATAAAGTGAAGCCTCACGTGATTAGTGGGTGGAACCAATGTAGCTTTTGAC  
GACCCCTTTGCTTTTACTCGTATCGTCAACATTTGAGTGATCACACCAAGACATGTCTTATTGTGTAGCAGATGCACTACAGCAGATCTGTCTCTC  
CTAGAGCAACAGAGGAGGAGGAGGAGGAGAACTCCATATAGATTGAGCACCCTCAAGAAAGAAATACAACTAGCAAGCTGGTATTTTCAATAAAT  
GGGAAAATTTGTAGACAAGAAACTGGCATGTTGAAACCTGAAATGACTCGAGATTTATTGGCCGGGCGAGAAAGTCAAGCCATACCAAGTTTAAGGAA  
CGCAACAAGTATCTCTCAGTAATAAAGGATCAGCAGGATGGTTCCGAGAAATTTATGGCGGTATGTGACAGTGTCTTCTGTTGGATCTCATGAAAGTGT  
GCGAAAGCCCTATAAGAAATCCCTCTCTGAATTTAATTTGAACGCTGCTCGCCAAAGTGAAGTGTGCTGCGGCGACAGGTTAAATGATGAAGATGA  
AGTAGACCTACACTTTTATCTATTGGGATTTCTGAAAGCTGAAGAGGCCAGGATCAGGCAAAAGTACACGCTATTGTTGCAAGGATGCTTACTTGACT  
GGTATAGTTTCTACCTCAATCAACAGGAAGGGGAGATTTTATGGCTGTGTATGGACTTCTGTTTAAACGAGCGGCTGTGACAGCAACCTGGCCACTC  
CTCTATGTATAGGAGGAAGCAATCTGTAGAAATATGGGAGAGAAAGGAGTGTGGGAGTGAAGAGACATGATTTTGGCAGATTTTGGCAGACAC  
AAAGGGAGGTATGGTGAGTCAACCTATCTCAATCATGTTCCCTATCAACAGATTGACATCAACAAGTTTGTACCCGATGACCATGTCTCAGAAATATCTG  
TGCACCACTACCTTTTGTGACCCATCGACAAATATTGCAACTGAGGAGATAGATTGGTACTTGAATAAATGAAAAACAAACACCGACTCTTTATTGTTGT  
TGGAGTTTATTGACGAGTGCAATCAGATTGTGTTGTCGAGTACAGACCTTGTATATTGACGTGCACTATGGAAGAAATAGCAACTCTAATAGACAAAC  
TCCAATTACTCGCATGAGGAAAGTTTGGGCTCTAAGATTCTAGAAATTTGGATGCGGAGAGACAAATATAAAACGTTGGTGCACCAATACATCTCCC  
AATATGAATGTCACGTGACGAGGTGGAATGTTTCTCCCGAGATTGTTGCTGACATTAATATGCACTTTTGGCGGCAAGGTGATGATGATGATGATGAT  
CCCCAGCAAGTTTAGAGTATATGCTTCAAGTATTGCCCAATATGTTAATCGACAGACCTACATTGGCGCACACATAACAGCTGGAATAATGCTGATATT  
GGAGGATATCTCTCAGAACTCGAAAAGGACTTTTCTGTTGAAAAGATGAGGAAATATAAGAACCCATTGGACATTTAAGGGTCAAAAACCAATACGAT  
TCTGTCTATAGTCTGTGACCGCCCAATCAGATTACAGAAATGCTGATGTTTACCTGAGGAAATATCCGTGATTATGAGGCAATGAGAAATCTCGAGAGTTG  
TTAGCTTATCAGACAGAAATTTATCGCCGCTTGGCGCATTTGATTGCGCCCAATGATCCAGCTGTTAGACTGTGGTCTTCTGCTCTAATCAATGTTGGAAT  
GTTGGTTAGGACATGGAACGTAACCACTTAAGGAAATCTCCCTCAATGCAAGCCACTTACAGAGCCGATCGAGTGTGATGTCAGACAAAG  
GCCAAGGAGTTTGGCAAGATGGGAGACATGAAAAGAGCTGGTCTTAAACAAAGTTGGCAAAATATTATGAAGCTCGGTATGAATTCATGTATGCTGATT  
TGGCCCTAAGAGCAGCTGTGAGCCGTAAGAGATTGCGCTGGAATGCTGCCAATACCTGCTCAAGTATTTCCAACATGTCAGCCACCGGAGGAAATGGAG  
AGGCAACAGGCACTCGGTGACCGCCCAATCAGATTACAGAAATGCTGATGTTTGGCAATATTGGTTGGGATACAGATGAGTCTTCTGCTGACT  
AAGCAGAGCTACGGGATACAGATTCTGATTCTCGCTGCATAATATTGTAGCTGATGAGGAAATGATACAGAAATATGATGAACAACTGCGCAATATT  
ATTATGTGATGGATATTGCTCTAAAAATAAAATGGCTGCAATATTCTCCATCTAGTCAACTCGTTAACAAGGGCATCCAGTTTGTAGAGCGCCGAGA

FIGURE 2, Sheet 39 of 40

GAGGAATTCGAGGCCCTATTGGTGCTCCTGGTAAGCCTGGCACGGAACGGGTTAGAGGTCCTAGAGGGGTGAGAGGTGTCTCTGGCTATCCTGGCGCACA  
 AGGGGAATTAGGTCCCCAAGGACCAACAGGTCTCAAGGGCCAGCAGGTCTCAAGGGCCGATGGGGCGTACAGGAGATACTGGTCCCATGGGCCCTCCT  
 GGAGCAGTGGGACCAAGAGGAGAGAAGGAGGTAGAGGAAGAAAGGGAAAAATGGCCCTAAAGGAGCGGACGGAAAGATGCCGTAAATATCATACAAA  
 AATATTCAATCACCCTATGCTCGTCAGAGATAATGTGGGAAGGAATGAAATCGGAGAAGCATACATTGGAAGATCTTATGGAAGTATGATCAATCCCTGT  
 GATGATAGAAAATAGAATAGGGATGACAAATGAGGACAAAAAACGAATATTGTATACAAGTAATGACAATGCACTCAATAACAACAGAGGAAGAACA  
 TCGGGTGTTTTTGTGGTAAGCAATAAGACAGATTATATCCTTTTAGTTACTTTACTGATGCCAGAAAGTGTTTCTGTAGAACAGATGTCAGTACAAATG  
 CGAGGTCAGAGAGGGTGAATGCTGTAGAGAAAGAGAAAGCAATCGTACAGATTTATTAGGCCGTCTGACCAATCTATAGGTACTCATTACAGTTCAAA  
 AATTGCCGTGGTAATGTATCCAGACGCAAGCATGAGTTACTCAGTTGATACATTAGACGCTGATGTGGCGCGAAGAGAAACAACGTCTGTGCTTTTATTA  
 GCAGAAACCATACACGGGAAAAAGATAGAGGTTTCTATGCTGATAGAGGAAGTGTAGGGAGGTGATGGTACCTCCCACTGAAGAAGAGTTATTGGTAT  
 TGCAAGC  
 (SEQ ID NO:1)

CT1035

Nucleotide

Genomic coordinates:

Start: 130589

Stop: 131444 (SEQ ID NO: 2)

Amino Acid

MAAAKMDAILADINGNDTDL SKLITDVIQKRAKAVMDRNRRAKMDMNRVDEAIQEAVA AK  
 KQKALVVFDKLV EETDSGQSVPTLSGSDYDAWVDRAMP SHIELVESVEGDSLYDKLP PF  
 NVQDIDDQIGDEIDTPISYLAMVVVKVDCETGDIEEEYNLAPTFGVTQNNKIYRDERDQI  
 FTKADKSVRIFKLAKLDSISGKSRQLTYAVKNNNEYTEFVCSVFAEFESDSDTTKSGIGI  
 REYDKPKNEFEYEEREIFTFFIPIQPAGTKLLLYFLVDVRSRII  
 (SEQ ID NO: 3)

Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q65326 (Q65326) DNA POLYMERASE (FRAGMENT)	36	0.31
MYSB_CAEL (P02566) MYOSIN HEAVY CHAIN B (MHC B)	34	1.6
O02244 (O02244) UNC-54 PROTEIN	34	1.6
O04327 (O04327) CELL DIVISION PROTEIN FTSH ISOLOG	33	2.1
Q04009 (Q04009) MYOSIN HEAVY CHAIN	33	2.1
O02077 (O02077) COSMID C48E7	33	2.7

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 765 of CT1035: this  
 corresponds to nucleotides 130649 to 131413 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=350

5'stop=373

3'start=405

3'stop=427

5'primer=TCCCTCCTTTTAACGTACAAGACA (residues 350 to 373 of SEQ ID NO: 2)

Tm5=57.95

3'primer=CCATGGCAAGGTAAGATATTGGT (residues 405 to 427 of SEQ ID NO: 2)

Tm3=57.97

probel=CGGTGATGAGATAGATAC (residues 387 to 404 of SEQ ID NO: 2)

probelstart=387

probelstop=404

direction1=Forward

Tm1=69.08

score1=1.91

length=78

CT1037

Nucleotide

Genomic coordinates:

Start: 131480

Stop: 132941 (SEQ ID NO:4)

Amino Acid

MFVISIATSLVLFVFFLLFVSITILDGAKTIDSQPFRRKRKRKRYRTSESGDGIDGGTGT  
 NGGGGGGGEGGGGGTNGNGTGTNGGGGGGEGGGGGTNGNGSGTNGGGGGGEGGGGGT  
 GGGNGNGGGNGNGNGGGDTDDFEPTALLKERLLNSISSKPKEYEAFVSAEVETAL  
 QLSRDDSTQTIIDDDQLELDASDTLQGKPRDYLKLAGVSSAFLEGTTRKAEDRARNI  
 NEEIEIAQTILSQLREKHINDEYDGKYATPEERADFSNSLNLVTKYTNHEVGLLVGETIEK  
 AFPHEIEFERCIILVEDFNSGTITSNTMQYRSNAYKIRVVEGSTTDPGEVVPDDCLVFAV  
 VVNKEQHSLEISATNRCQDICFVIIPRLSAIGKNATMVIRKGEIKQETYLEVANKNDTT  
 HFSIITDKDES VGIELNMLIFSERILPTLSDPATVPRPLTDANVLSAYGKRLGVGAFTDK  
 NLLSSQ

(SEQ ID NO:5)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O65514 (O65514) PUTATIVE GLYCINE-RICH CELL WALL PROTEIN	120	3e-26
Q21835 (Q21835) R08B4.1 PROTEIN	112	6e-24
Q43522 (Q43522) TFM5 GENE	110	2e-23
O53553 (O53553) PGRS-FAMILY PROTEIN	108	1e-22
O65450 (O65450) GLYCINE-RICH PROTEIN	107	2e-22
GRP1_PHAVU (P10495) GLYCINE-RICH CELL WALL STRUCTURAL PROTE	106	3e-22

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 392 to 1365 of CT1037: this corresponds to nucleotides 131892 to 132865 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=723

5'stop=745

3'start=792

3'stop=811

5'primer=TGAGGAAGAAATTGCACAAACAA (residues 723 to 745 of SEQ ID NO: 4)

Tm5=58.39

3'primer=CCTCGGGTGTGGCATATTTT (residues 792 to 811 of SEQ ID NO: 4)

Tm3=58.37

probel=CAACGATGAATACGATGG (residues 774 to 791 of SEQ ID NO: 4)

probelstart=774

probelstop=791

direction1=Reverse

Tm1=68.98

score1=1.98

length=89

WO 01/38351

44/201

PCT/US00/28888

CT1038  
Nucleotide  
Genomic coordinates:  
Start: 132993  
Stop: 133896 (SEQ ID NO: 6)

Amino Acid  
MGDKQKVEQLLRELKAEANDDWLSANVDPIVERFVTTKSDETAQVVVKQAVDEKYDELLED  
KVEEMRPDIINEASETYDKLAADMIREVDTSSVIAPAIAGTVARTINNLRDKRKEYEKRL  
WTLAYKPWRRYVQAITVMEFRLSYKD LTVHANS DTYLTFPFLRIKKIAYINNDRASPVNC  
SLSVSYPNKSEWGNDNGVGRKVDIHIRRNDLQEKDLYLSVICMLD TDFSGYDKAVEVDAH  
KFHFEAGNRMTMFLPKTSNLFNRSHIVNSKICTIVEFPASASSASTTELDNVYYRITCTCS  
(SEQ ID NO: 7)

# Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
RESA_PLAFP (Q26005) RING-INFECTED ERYTHROCYTE SURFACE ANTIG	38	0.12
Q9X9G6 (Q9X9G6) HYPOTHETICAL 35.5 KD PROTEIN	37	0.15
CAA21397 (CAA21397) ORF 74	37	0.15
Q9ZFZ9 (Q9ZFZ9) HYPOTHETICAL 21.1 KD PROTEIN	37	0.15
RESA_PLAFP (P13830) RING-INFECTED ERYTHROCYTE SURFACE ANTIG	36	0.26
Q9ZH03 (Q9ZH03) LAMBDA HOST SPECIFICITY PROTEIN J	36	0.45

CT1039  
Nucleotide  
Genomic coordinates:  
Start: 133968  
Stop: 136344 (SEQ ID NO: 8)

Amino Acid  
MEEESQVRVQRRIGVLPEEAASQILKDTKLRSYLVGVHGWGYSVSVIKSALQKGCRRNDED  
ITAWSIREAYLYYHLGLNYIENVKPAAKSLNTNMVNRIKIIAVEDTSPRSMVASNECVRT  
LEKYEKGNFRQPSYLMDAAMRLVHASSSRVCSHMRALCCKEEDSDKLGGIYYANFNELET  
QCVSAVNFSPIERIKHVFREIESVVLGKKSQVLLNLRSAAYHVLRYYGDKVKDTNKKHS  
GPFKRKEFEQFWGLCFKFVTQHVKTDPRLCYFNELTYAINWRRDFFCSKGGFFREESLFL  
TSIVELIIAMCIGDRKQFAKIQKRDLLKRFNKGEGRKEEAATFDWIEGHVVRMPQMPVWV  
LDKHTNKNTHGVSFALSSMVSGGDKRWSPGVWLHSYTKMRLDSPPPPEVGGQFLDQAFNT  
LKREAASHCVTRNICTTTGFIKASSFTANINSEPMEIKEEIKKKRIEIKDDNTTATVTVS  
ATTSSSITSTPPPTKKQKTTTPSGSNKVDSIQLNNLPTLNMEDLDRVLEVHNQNSKKGVAA  
TVLMKDGKNKVVFKEMRKSFGWGSQNFVQVLKDEDVCKLDYLLPCPDSPYRGLYRCYFK  
IVKDEISSTAARIEKVKGWGENAMCYFISGCVTRQEGIGKIITDVRLSHMGPNKQYVYDNY  
RQLIHILIFRLLTGVSNTNTSNILVGDGGNLFSVDENYVGAKDPRTALENRKIKELQLLL  
KTSFKVNKVKEDIDSCLPWLFDTSKSDKIMNGVCNIGKMGIGPTTLDIVKNNCTCIL  
GVVNDLLYDNK  
(SEQ ID NO: 9)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q9Y1H4 (Q9Y1H4) GAG	43	0.011
RAP1_YEAST (P11938) DNA-BINDING PROTEIN RAP1 (SBF-E) (REPRE	41	0.031
SLY1_YEAST (P22213) SLY1 PROTEIN	40	0.070
P90603 (P90603) MUC.CL-1	38	0.20
O52224 (O52224) GLUCOSYLTRANSFERASE (EC 2.4.1.5) (DEXTRANSU	38	0.35
Q20202 (Q20202) F40E10.5 PROTEIN	38	0.35



## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 1092 of CT1039: this corresponds to nucleotides 133866 to 134957 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=1255

5'stop=1277

3'start=1353

3'stop=1375

5'primer=AATACATTGAAGCGAGAAGCTGC (residues 1255 to 1277 of SEQ ID NO: 8)

Tm5=58.71

3'primer=CCTTAATTTCCATCGGTTCAGAA (residues 1353 to 1375 of SEQ ID NO: 8)

Tm3=58.18

probe1=TAGTCATTGCGTAACGAG (residues 1278 to 1295 of SEQ ID NO: 8)

probe1start=1278

probe1stop=1295

direction1=Reverse

Tm1=69.03

score1=1.96

length=121

CT630

Nucleotide

Genomic coordinates:

Start: 249215

Stop: 247358 (SEQ ID NO: 10)

Amino Acid

MALKDAFTERLVVNKVGSGTDMAPVVEDDRQKSLFQKVENLYRVLVVEQKNSAITLSGNK  
 NTNKRQCRQVEEDKVFEGEDRTVSNLPQAVKETIAANAESILDYWKNVIPLLDTKKER  
 SGKSDTFLRTAVICLVRCVSYKDMKTCSLIYEFHKILNKSTLDPLLKDILDNKQELLH  
 MDSKYGSKTTSPELAKETIEALYTTVYNHWTNAFKLYQASLTHKPVTGKKYASVIHFIRT  
 WRKIVKAYVSKHNNVERDLSLKNIMKNESADNANVLTIEKMYKKIGNSVKNTNNNSAHQM  
 SDESEDDDDDDDDCEGMDVCDEASEREKKHQESLYPINTPVTITGDYIFKVLLELVLS  
 HIHPWKIPMCDFVNRNIPKLMKAMETDISNAVIEVRASKVNPVQILPIAANFWDCKSG  
 KPPSDVKFCMMFNEPSSNETLSSGAGVFGFRFIGGPFSSHKSKELDIISNCLRSLLLNEAD  
 NLSTRIWREGGSVVCFNYPITARGAVLGYGEQLSERSIKALWAKKIQDAVTESVKRQRN  
 AADKNSRNCDDLGGDEGVVSMKTVTFGCANMLKTQNGMGKFNVVVSFEDSIQANKEGAARQ  
 YMSQQVFTHSFPALDQGK  
 (SEQ ID NO: 11)

Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
O97318 (O97318) PFC0865W PROTEIN	45	0.002
YB00_YEAST (P38114) PUTATIVE 126.9 KD TRANSCRIPTIONAL REGUL	44	0.003
Q08281 (Q08281) CHROMOSOME XV READING FRAME ORF YOL138C	42	0.010
Q92271 (Q92271) 12.8 KBP FRAGMENT OF THE LEFT ARM OF CHROMO	42	0.010
O62235 (O62235) F36F2.3 PROTEIN	41	0.018
Q83970 (Q83970) (CPV)	41	0.023

CT631  
Nucleotide  
Genomic coordinates:  
Start: 264975  
Stop: 259164 (SEQ ID NO: 12)

Amino Acid  
MGSKRPCSSGQEPVTKKQKNNNNNSNPVPVINIKSYPFLLATRTQVLRSAVAAAAASPSG  
SSSSSSSSASAVKLPDTCKEARKVLSTVSLQQSLAVRYLCNSISVSYAGGGISVFHLGG  
PGAGKTTMVKELIAVLNDHGLIDSGSADMLLCKCKNSAKESLMCAKCKPGGSSLMYPESV  
FSTLNKGFEIPVIFRKDEITLEKIQFVADKLKWKVIQVLNLRFLVIDEYTMASCRELVF  
IDAVLRIAKHRPDIPFGGVFVILLGDNQRNSAVVEDNTNHIQKKIKNPSEEEKPKNNKN  
NKNKKKKKEKKEKGEEEGDENEEEEEEEEEEEEESDDEAETKKEEEKSTFFQGSVEQD  
NFGQEDNAKLYTEVFIKILKMFCSDFFGNPSNLRNIVNKRHEAILMKSNNVKSNNNLV  
SSAIKVEDCGNASNKKEVTAPSSSPAQSTAEENCDEFDDEEDDDFFNNEAFLKLMERNAL  
EKDRASGALNGFSLRCKSISDANEKIRSGTTSVSDKSSLDMMKSLPLSALIEEGICSEL  
AHISELKMSNANLEKYTENVCISVFDMMAKAMREIDYSGREKLYIVSSLSERFKDTHLT  
SLMDEEILNVKYVHGSDPKCIDAVPFNSAHNRAVAACVRNAFFRDGKDFVDETPIAN  
FKDNLRTVASFLENETLTYKELLAKSENIRSILLKKTGNNSASSRTAAAAAYEDED  
YCYFDEEEAMDLEDGGSGGSGMKSSGGGGDDDDDEESGEMIYRTDIPDKLHRDASTLDRVG  
HLVDFHVVWKKWLTENKPSDLVRARVWYLYTLVRMQQVKFDNGKLPKSLDLSALSGRLFHS  
PSEWATSTGVGVGGGGGAGADKPLHDEYWLRLVLSMPISTGGDVGKSMILLPAYSSYLSALS  
RTYIMSSLKRIDIKHAYSLMYGISLFDMTANLQDLVDTRMAGRSSRNGSVFMDNFDPVQ  
YFDNIFPSMVNEFLMYRKEDVFNNQMMEGVKGLKISRVLQTAHTENNNINNRHNSLK  
YSEKSIVLAMQMVTSISKGNERRKIEEFITKEQQQPKDMCERLMANSKAKQEKDAISSK  
TDRMMGAILTLTKKHVLKNAVSNLVDTSIIKETKNNNNSSSSSTSLAAAAAVENSVPAL  
RVEVKFVNLNMDLSDISHEKTI SHKYRQQLINAIKTRSTPLFDKFTDRKILRAAESPRAL  
TTILLDEKKKVTRAKSITLYQQQNVIFTTSNRMIHGTQERFVTKDTGVVTNLMYKNGELT  
VFVYVERLGQKCLEIKEGRQIIIGNPNIKNGGFGNNVYVQYLPFESSQAMTIYSCQGHTEFF  
RDTIVDLSGASTQDAYVAVTRNSNPQNLFIIONHSVERGNLCNIKAMSKDKAYTMPFIGG  
IADENGSDFINHDTVSVSREVAESSAAMDDDDYNGDGGVTMYSAYDPSKDVVAAAEFFILS  
RSGKLSLFSNASWMANTAKVIOQHGLETELKNIRDFFFGVNNGDVAKHYEKLCKNKKMIELY  
TAIVRSITHYSIASGIVKQPSSKLCEYETKQKNKKDYIKIHPVFVNRAPKESTIEMLLF  
DIAPHNKATIVFQFYVHYIFLVYEKLNVLNSSFAFLPSPNPNCLNQYVRPKSITNSTHVP  
NLGYESKDFAHCKDGGERDVKLRPLITSADFEFSNNIEGILKKVSDTSNQNKVNKYMDVVC  
KSMQHNLRRTGKFCRPTETCGLSKHGSIVTSTCTAQEKGENIHVDAEKGLCMSDEANVY  
CMLMFMSKIAAASGVSEFPKIDKSILESNPETPSDTISLLAPRKTISPTNNLHFSMSDEV  
LFCCQVHMPMKRVQFSLHVKRTGGALKSTFEEEGLPTKIFSPNFATYPLFKCKMYGAI  
IAMTEMQGHFAKYSTLDIRKSMFTGVGTVDLEKISGEGNEVMDKVDKFIVKNVSNILF  
KEQGKRVSVFFVSCAIH  
(SEQ ID NO: 13)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q9YTL7 (Q9YTL7) ORF 48	68	5e-10
Q35788 (Q35788) CYCLIC NUCLEOTIDE-GATED CHANNEL BETA SUBUNI	68	5e-10
NAB3_YEAST (P38996) NUCLEAR POLYADENYLATED RNA-BINDING PROT	68	8e-10
Q07034 (Q07034) RNA BINDING PROTEIN	68	8e-10
O96229 (O96229) HYPOTHETICAL 78.6 KD PROTEIN	66	2e-09
O96134 (O96134) SER/THR PROTEIN KINASE	66	3e-09

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has sequence identity to nucleotides 1 to 830 of CT631: this corresponds to nucleotides 259304 to 260133 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=2408

5'stop=2429

3'start=2489  
3'stop=2509  
5'primer=GAGCCCGTGTGTGGTATTTGTA (residues 2408 to 2429 of SEQ ID NO: 12)  
Tm5=58.49  
3'primer=AGCGTCCAGACAGTGCAGATT (residues 2489 to 2509 of SEQ ID NO: 12)  
Tm3=58.56  
probel=CACTAGTGAGAATGCAAC (residues 2432 to 2449 of SEQ ID NO: 12)  
probelstart=2432  
probelstop=2449  
direction1=Forward  
Tm1=68.99  
score1=1.99  
length=102

CT632

Nucleotide

Genomic coordinates:

Start: 268485

Stop: 267717 (SEQ ID NO: 14)

Amino Acid

MAGVDLYGGHIKPYGETVFNNKMQGNRGKIRALINEKAAATLPMSEDNISAWVTEVAADV  
 FPDPKSALTFFVPNKSLNAFAWDVLKTPASVEIDIGKRIPQLIENLHMSDFTVAIFRVKC  
 DDQGRYETSYNLSFSMGGKINHGLIRTLAKAQDIVVWKRDFSLTIENFEVDNGKKRLDFL  
 FNNQTDKSCFVKIFHEMESEKDIAIKPEKRGSSAVWDEVYSDIVTKNTRNAKFSRLRYRNE  
 KPVDHLLLYCMVTYF  
 (SEQ ID NO: 15)

Top Blast Hits

Sequences producing significant alignments:

Score	E
(bits)	Value

097036 (O97036) PLC-BETA2	34	1.5
045329 (O45329) F09C6.2 PROTEIN	33	2.5
P97868 (P97868) PROLIFERATION POTENTIAL-RELATED PROTEIN	32	5.6
P70287 (P70287) RETINOBLASTOMA BINDING PROTEIN 6 (PACT) (FR	32	5.6
AAD49229 (AAD49229) EHEC FACTOR FOR ADHERENCE	31	7.4
CAB55629 (CAB55629) LYMPHOSTATIN	31	7.4

Comments:

TaqMan Primer/Probe Sets:

5'start=378

5'stop=403

3'start=452

3'stop=477

5'primer=TGAAACCAGCTACAATTTATCTCCTT (residues 378 to 403 of SEQ ID NO: 14)

Tm5=57.68

3'primer=CCTCTTCCAGACTACAATATCTTGGG (residues 452 to 477 of SEQ ID NO: 14)

Tm3=59.43

probel=TCAGAACACTGGCTAAGG (residues 434 to 451 of SEQ ID NO: 14)

probelstart=434

probelstop=451

direction1=Forward

Tm1=69.15

score1=1.84

length=100

WO 01/38351

51/201

PCT/US00/28888

CT633

Nucleotide

Genomic coordinates:

Start: 266790

Stop: 266442 (SEQ ID NO: 16)

Amino Acid

MASPLVASLGGGKNILFGLLLITIIIVIVIAVIIIKAPLLASLLAGTALAGTIASALGSIP

GVGGAFKKAFGKKGKGGPKTPDGGAKKTNQKPKKGKKKPPTRRSIFKRIPKIKF

(SEQ ID NO: 17)

Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q02391 (Q02391) CYSTEINE-RICH FIBROBLAST GROWTH FACTOR RECE	36	0.092
Q91019 (Q91019) MUTANT CYSTEINE-RICH FGF RECEPTOR	36	0.092
Q69526 (Q69526) GLYCOPROTEIN B	36	0.12
Q9XZ15 (Q9XZ15) HYPOTHETICAL 29.3 KD PROTEIN	36	0.12
Q14113 (Q14113) AORTIC CARBOXYPEPTIDASE-LIKE PROTEIN ACLP (	35	0.21
P79922 (P79922) MODIFICATION METHYLASE (EC 2.1.1.73) (CYTOS	35	0.21

CT634

Nucleotide

Genomic coordinates:

Start: 283360

Stop: 282673(SEQ ID NO: 18)

Amino Acid

MVSSRTSTTSSSAVAATSTLLPTKRKREPEEVKVKVEVKMEQEELVEDSSSNKRPRIKEE  
 KEEEHKETHHLSLPCKEEEDDGEYEEYEDRVDDDTAEKMENLLVQLDNTTK  
 NIKLKNPLREHDMVSHYEHFEVQNTVNFSGVLSDIGFLINREAVSRWGNTPPPKEFG  
 DMEIGSLTVNQHLHKCDNFVQAVVQVKVEDITPSIEVTIDSLIDDPW  
 (SEQ ID NO: 19)

Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O35788 (O35788) CYCLIC NUCLEOTIDE-GATED CHANNEL BETA SUBUNI	60	2e-08
GARP_PLAFF (P13816) GLUTAMIC ACID-RICH PROTEIN PRECURSOR	57	8e-08
Q9YTL7 (Q9YTL7) ORF 48	56	2e-07
O94922 (O94922) KIAA0835 PROTEIN	55	3e-07
O08995 (O08995) MYELIN TRANSCRIPTION FACTOR 1	55	3e-07
Q93424 (Q93424) HYPOTHETICAL GLYCINE-RICH 37.0 KD PROTEIN E	55	5e-07

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 575 to 1 of CT634: this  
 corresponds to nucleotides 282678 to 283252 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=210

5'stop=231

3'start=287

3'stop=308

5'primer=CCTCTCCCTCCCATGTAAAGAA (residues 210 to 231 of SEQ ID NO: 18)

Tm5=58.19

3'primer=TCGTCGTCCACTCTGTCTTCAT (residues 287 to 308 of SEQ ID NO: 18)

Tm3=58.40

probel=AAGAGGAGGATGAGGAAG (residues 266 to 283 of SEQ ID NO: 18)

probelstart=266

probelStop=283

direction1=Reverse

Tm1=68.99

score1=1.98

length=99

CT635  
Nucleotide  
Genomic coordinates:  
Start: 285773  
Stop: 284075 (SEQ ID NO: 20)

Amino Acid  
MARSVGLLSVTPEYDTFKYIKMEEFKTLKVKNFTISGENPDKYEHILLSFKSVDRVTKS  
ELRDGLYIVRLKDKVHLHIKNGVHRLRQLTGDNTLQVGLKYTHNLPRLGSLQDDGCEDY  
GEKWNESLPIDMQNINKIVKEKALLSDKNFKFSPLYRLLHERLSNAAVKKCDYMIITDF  
LVGCGYTPSHCPRTLRLNMEQLLVEQCGFSSRISVYDICDRLTYKGAYIANPITGSYSNMC  
LIVPMDKLGILIFYNSTHPSAKSIGNYMSSLFNATVIYANERDNLQMDNFRREIKFAENEV  
NMKEEELKELRKRCVSEEQRISLRDVHKKSSIATSRVDGACLVFAFSRDRDFSLLCRTN  
NGSFYSATEEGIRYVSSPEYKKRDVGERRPRLIMSITGSDAPICIRDSVRNHFKTRLFS  
RTSGNSITFAVPPGERELMEMVREVTGTDIKIFMDNGKVYQNGAEINVIDPTSKEYKELL  
KREENLPEDERKRLRRERRMIFNTSRAISMYNEERGDGSGGETSEDGDGNGSTSSKGEK  
RKREENEGNEYVLLNKACKDIKVC  
(SEQ ID NO: 21)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q21885 (Q21885) COSMID R09H3	37	0.32
O77336 (O77336) PFC0425W PROTEIN	37	0.42
O00164 (O00164) RIBOSOMAL RNA UPSTREAM BINDING TRANSCRIPTIO	36	0.71
UBF1_HUMAN (P17480) NUCLEOLAR TRANSCRIPTION FACTOR 1 (UPSTR	36	0.71
YPT2_CAEEL (P41880) HYPOTHETICAL 21.6 KD PROTEIN F37A4.2 IN	36	0.93
YG2K_YEAST (P53253) HYPOTHETICAL 106.5 KD PROTEIN IN CTT1-P	35	1.2

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 501 to 1 of CT635: this  
corresponds to nucleotides 284047 to 284547 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=686  
5'stop=706  
3'start=799  
3'stop=822  
5'primer=CAAACCCAATCACAGGCAGTT (residues 686 to 706 of SEQ ID NO: 20)  
Tm5=58.55  
3'primer=GGTTGCATTGAAAAGAGATGACAT (residues 799 to 822 of SEQ ID NO: 20)  
Tm3=58.19  
probel=ACTCCAACATGTGCCTAA (residues 707 to 724 of SEQ ID NO: 20)  
probelstart=707  
probelstop=724  
direction1=Forward  
Tm1=68.89  
score1=1.89  
length=137



CT636  
 Nucleotide  
 Genomic coordinates:  
 Start: 286706  
 Stop: 286076 (SEQ ID NO: 22)

Amino Acid  
 MIVFVEGSPLTGKTSWVDNMRTAGKGKQSFLNFMYTNRYDYLPIFPWTIQEHLRASDYQE  
 RPRLVDGMFGSSLNFFTGMWRHDTEQFPESKIGLREYLEMYGEEFKACVAEWVKYKPVFH  
 VMVYREEDVKMEPIIQELNDAHNWFIDVLKEERALFVKIEVIPRNVYKGNICSSCFSTS  
 KNYVYRVGKCTNSIVHCDMKCKFIAEKII  
 (SEQ ID NO: 23)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
P87968 (P87968) ENVELOPE GLYCOPROTEIN, V1-V5 REGION (FRAGME	33	2.0
P88315 (P88315) ENVELOPE GLYCOPROTEIN (FRAGMENT)	33	2.0
P88314 (P88314) ENVELOPE GLYCOPROTEIN (FRAGMENT)	32	2.7
DYHC PARTE (Q27171) DYNEIN HEAVY CHAIN, CYTOSOLIC (DYHC)	32	4.6
O40068 (O40068) GP120 (FRAGMENT)	31	6.0
O40067 (O40067) GP120 (FRAGMENT)	31	6.0

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 594 of CT636: this corresponds to nucleotides 286111 to 286704 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=239  
 5'stop=259  
 3'start=317  
 3'stop=335  
 5'primer=GGAGGCACGACACAGAACAGT (residues 239 to 259 of SEQ ID NO: 22)  
 Tm5=58.81  
 3'primer=CACTCGGCGACACATGCTT (residues 317 to 335 of SEQ ID NO: 22)  
 Tm3=59.42  
 probel=TGGAGATGTATGGAGAAG (residues 293 to 310 of SEQ ID NO: 22)  
 probelstart=293  
 probelstop=310  
 direction1=Reverse  
 Tm1=68.95  
 score1=1.95  
 length=97

CT637  
Nucleotide  
Genomic coordinates:  
Start: 300432  
Stop: 299085 (SEQ ID NO: 24)

Amino Acid  
MGGEDSFDDRYDSDALWENEGAKSIQVKETDLEVYRMHRAVPTLEEKNR TALRYYS DWS  
PVYRVPLFSLKDGSDPHERDFSFNVDPRRFGKVPVKVRRVDVRNPSRTAAIFVPTGPGLH  
VSSYTG DGM LVC PNHNFIGDLCSEIASDITIYNTSSSGRLSYATNFNSVEDNSPVGILFE  
TLPDDKM FQQVSIFSATEPASNISIGPM SHVKIKLGY YDEENATAVG VIRYGG LFYTSVG  
ACI IPEGVFFDDVVG NHSSMNIYNMTNQPK EIVLKEPRGEDAMEEDDGEEADYNFLGYVV  
RFEHDLKM QAMSSAYSSVSIDINSSSFHKCFLIKPKYNSILQPLVSSEVV LNDLSL NTRG  
REVEFHDRLP SG AQDNSYSIVKYM KIVSLKEGLKVVNPIINTELYKKKQALKVHVLNMTR  
DVGGLDTSEHSFGVIVCHAAKLPEVIGQ  
(SEQ ID NO: 25)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O23592 (O23592) CARBOXYL-TERMINAL PROTEINASE HOMOLOG	34	1.6
YGA1_YEAST (P53199) PUTATIVE 3 BETA-HYDROXYSTEROID DEHYDROG	34	2.1
VILI_DICDI (P36418) PROTOVILLIN (100 KD ACTIN-BINDING PROTE	33	4.8
Q925B9 (Q925B9) PUTATIVE TRANSFERASE	32	6.3
Q9YUY4 (Q9YUY4) ENVELOPE GLYCOPROTEIN	32	8.2

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 3 to 1254 of CT637: this  
corresponds to nucleotides 299132 to 300383 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=678  
5'stop=702  
3'start=788  
3'stop=812  
5'primer=CGGTGTCATTAGATATGGAGGATTA (residues 678 to 702 of SEQ ID NO: 24)  
Tm5=57.26  
3'primer=TCTTTTGGTTGGTTAGTCATGTTGT (residues 788 to 812 of SEQ ID NO: 24)  
Tm3=57.75  
probel=CTACACCTCTGTCGGTGC (residues 705 to 722 of SEQ ID NO: 24)  
probelstart=705  
probelstop=722  
direction1=Forward  
Tm1=68.85  
score1=1.85  
length=135

CT1040  
Nucleotide  
Genomic coordinates:  
Start: 137588  
Stop: 139940 (SEQ ID NO: 26)

Amino Acid  
MAAAVSGEGRISADLLLLLEQLTPDGDVIRYDSEQYTKPRKIFGDKSVIETIGHFLIHNH  
NQGESYQIASSVLEKFPALLNCIWNGESGGMALWKALYRAKKYRLNSLLVHKIKNWPSV  
AVIPIYGSVCDREERPIIMSEIIDKETLQTICKSDIRSLGMMNAKHGTLGGNLFHFYAR  
STKPFENFQYEAMGANAVLMAAEAIYDGRDHGLNPSEYTFPGLESADVGNPVEIAIS  
GDDDNMLNLNICNYGVSYEKTRGRVNRSLDLFLKMNTASKCLSVLKFEKHFKIESNTPK  
GEFEEKAETCVNCLDRNNVLTGSEQESYKLSGHLVHVKCLRNICIVSQHLRCEKCLKR  
FDESILRKCTPNLNWWTMPAGAGNEEEICFMRNKKLVDDFRKLLSPVSI PHFFKNSRQR  
NLDMLCPYSDHTIIPNKEDPKKNEDGNRVRVNHTAISEKQNKKEEDARIKRVAVRTFTAI  
REKQNKKEEDARIKRAVDMAVAAINEKNKEEDARIKRAVDMAVAAINENNKEEDARIK  
RAVDMAVAAINEKNKEEDARIKRAVDMAVAAINENNKEEDARIKRAVDMAVAAINENN  
KEEDARIKRAVDMAVAAATNEKNKEEDARIKRIIDLTVDMRIQRIVDMAIAAATKKDKK  
EEEKRTKREQELRADLRRAMD MVNEVQKKLEDMELEKGCNKDEAKNTSNVVSSSSVVAYS  
KEIVPCLGNNNNNAVIGMTSTNYSANNTKNNVFGSPHKFSFNDASRFSNIVETPKMSFNFS  
FKT  
(SEQ ID NO: 27)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q9ZU69 (Q9ZU69) PUTATIVE VICILIN STORAGE PROTEIN (GLOBULIN-	71	4e-11
Q26775 (Q26775) TB-292 MEMBRANE ASSOCIATED PROTEIN	70	5e-11
INCE_CHICK (P53352) INNER CENTROMERE PROTEIN (INCENP)	67	5e-10
YDF3_SCHPO (Q10475) PROBABLE EUKARYOTIC INITIATION FACTOR C	66	7e-10
Q26774 (Q26774) TB-291 MEMBRANE ASSOCIATED PROTEIN	66	9e-10
P91257 (P91257) SIMILAR TO C. ELEGANS UNC-89	65	2e-09

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 853 to 415 of CT1040: this  
corresponds to nucleotides 138912 to 139349 of the genomic reference sequence.

CT1041

Nucleotide

Genomic coordinates:

Start: 140110

Stop: 141616 (SEQ ID NO: 28)

Amino Acid

MVYKGFVSPSFITIRMTSNRPTTSPLSFSEGFSLSGDKYDITYEDILLEQFNCFKTSSPS  
 SARKSEIEDKTLIFQLKEGEKFLAKGIEELREILDDNSATIEPIISPTTFNDRNELLNH  
 EGDISSSPLYTQIMKHISPEHDIYELDLIVGTDLLFGLGVNLRNVSKLMKKISYGT LNVV  
 DVCHRKFFNNRIIVNPISSSFKNVCIIPLFSAAEFSSSLGECRDLFNGICDDVERYINS  
 YFFYPENTTTTTTAPSSPEMEIADEEEQSPKTIKRNDNASRNWSGVCLIFEVFKNTYYI  
 INRGDRGGSFEKAVKSAISSIKEKRCKITDINGNKPRLMVITGCTELYFKDALKQIGE  
 NRRKFLKMNGNYFLIDEQADLIEFAMSVSGAGERIFVNGLGMFQNRKMIPVIDPLTYEN  
 VVCGEHDIQKEDAILSVRRAIADYND FVSKNKRGGKRSAAEENEDEDADASSSSSSSPPP  
 SSPPAHKKSRLPDEGEKCTLC  
 (SEQ ID NO: 29)

Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q26258 (Q26258) BR2.2=BALBIANI RING (5' REGION, REPEAT UNIT	38	0.12
Q23804 (Q23804) SPID PRECURSOR (FRAGMENT)	36	0.82
O97324 (O97324) MAL3P8.1 PROTEIN	36	0.82
Q99112 (Q99112) HOMEODOMAIN PROTEIN BW2	35	1.1
Q99111 (Q99111) HOMEODOMAIN PROTEIN BW2 (FRAGMENT)	35	1.1
YA55_METJA (Q58455) HYPOTHETICAL PROTEIN MJ1055	35	1.1

Comments:

TaqMan Primer/Probe Sets:

5'start=822

5'stop=846

3'start=929

3'stop=950

5'primer=AAAGAGAAATGACAACGCAAGTAGA (residues 822 to 846 of SEQ ID NO: 28)

Tm5=57.46

3'primer=GCACTCTTCACAGCCTTTTCAA (residues 929 to 950 of SEQ ID NO: 28)

Tm3=58.24

probel=AACTGGTCTGGTGTCTGT (residues 847 to 864 of SEQ ID NO: 28)

probelstart=847

probelstop=864

direction1=Forward

Tm1=69.07

score1=1.92

length=129

CT1042

Nucleotide

Genomic coordinates:

Start: 141695

Stop: 142541 (SEQ ID NO: 30)

Amino Acid

MAVNLDNVILVNINNKDEDLTKLVSEAIKRRRAKTVFDTKNQAGFDMRRQVEAALYEAISSKKK  
 EKAIAKFDELIQERGDEITPLTTMQYEEWVNRTITPSLTENLLGDVEHADFLDRMTPVS  
 EEDIEGFAASTFKEVSDSKTATVIVKADCETGDIDEVYNLAPSGVTQEIKIYRSNNSSEL  
 DNVADSFHIIKISATDSGNTKKLLYGLRNKKAGYTCLCRIFAEIESDGIMANTNIGVAE  
 NNRDEIDENEEGKYGFLIPKQFAGAKLIYYFFLNCWTX  
 (SEQ ID NO: 31)

Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
gi 2492980 sp Q10043 YRP1_CAEEL HYPOTHETICAL 37.6 KD PROTEIN R1...	33	0.79
gi 401691 sp Q00933 YSCI_YERPS YOP PROTEINS TRANSLOCATION PROTE...	32	1.8
gi 1706480 sp P51892 DNL1_XENLA DNA LIGASE I (POLYDEOXYRIBONUCL...	32	1.8
gi 267570 sp Q01250 YSCI_YEREN YOP PROTEINS TRANSLOCATION PROTE...	31	2.3
gi 3121979 sp O07597 DAAA_BACSU D-ALANINE AMINOTRANSFERASE (D-A...	31	2.3
gi 6686325 sp P71018 PLSX_BACSU FATTY ACID/PHOSPHOLIPID SYNTHES...	31	4.0

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 700 to 1 of CT1042: this corresponds to nucleotides 141748 to 142447 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=350

5'stop=367

3'start=392

3'stop=414

5'primer=GGACCGAATGACACCCGT (residues 350 to 367 of SEQ ID NO: 30)

Tm5=57.70

3'primer=CCTCCTTAAAAGTAGAAGCAGCG (residues 392 to 414 of SEQ ID NO: 30)

Tm3=57.66

probel=AAGCGAGGAAGATATTGA (residues 368 to 385 of SEQ ID NO: 30)

probelstart=368

probelstop=385

direction1=Forward

Tm1=69.01

score1=1.88

length=65

CT1043  
Nucleotide  
Genomic coordinates:  
Start: 142610  
Stop: 143699 (SEQ ID NO: 32)

Amino Acid  
MTVLAVYTAPQIKSKKRKIEDENEEPVKTLEDFVKGRLLNAVKEKPAEYFELLISADT  
EAALKTAETALRDFVIENDSVEIDVEEVLEEKPREYVFKLAGATSETLTNTIIAEVQKK  
AALITEEDITIKMLKQFRAANKDNKDGEATPEEKEDFTNNSDLVGLYLNEVVEKTTNIVI  
NKIFPHEMVFERCAILIEDFDTGVVTDQAIQIPSNKYKIRLVEGDEPEVFPDCLDLAVS  
VDKINHVLKISAKNGCENNCFVLIIPRFSPVGSVSSMILGSTDQVKPKTFLFLANKNDSTH  
FQFTMDKQHSVGCELDMLIFSERNLRLNLPDSKPRPLSDADILASYGKRLGTGVFTTENLV  
DD  
(SEQ ID NO: 33)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q26938 (Q26938) KINETOPLAST-ASSOCIATED PROTEIN (KAP)	40	0.022
Q23332 (Q23332) CENTROMERE PROTEIN HOMOLOG	39	0.037
Q9ZES5 (Q9ZES5) CTC PROTEIN	39	0.049
AAD45753 (AAD45753) ANON1A3 (FRAGMENT)	38	0.083
O66878 (O66878) CHROMOSOME ASSEMBLY PROTEIN HOMOLOG	38	0.11
AAD45759 (AAD45759) ANON1A3 (FRAGMENT)	38	0.11

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has sequence identity to nucleotides 1 to 982 of CT1043: this corresponds to nucleotides 142640 to 143621 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=543  
5'stop=567  
3'start=635  
3'stop=659  
5'primer=CAAAATATTCCCTCATGAGATGGTT (residues 543 to 567 of SEQ ID NO: 32)  
Tm5=58.37  
3'primer=CTGATTTTGTATTGTTGGAGGGTA (residues 635 to 659 of SEQ ID NO: 32)  
Tm3=57.33  
probe1=GACTGATCAAGCCATTCA (residues 615 to 632 of SEQ ID NO: 32)  
probe1start=615  
probe1stop=632  
direction1=Reverse  
Tm1=69.01  
score1=1.98  
length=117

CT1044  
Nucleotide  
Genomic coordinates:  
Start: 143759  
Stop: 144689 (SEQ ID NO: 34)

Amino Acid  
MSSSSSETPKTSTDTGEERIKDIVNALDNNGEWLSSYIDPIINNHISRKTAETVQKINQE  
VDERYDRKIADKINEIKSSIFTSAQTMVDQYAITFQEGKGANGTGPVMPVNTVIDTTL  
NKMGRNMLEYAEDMWDGDDWKRFSSTMTLEFDLSYSDLTMMRGSDGYFAFPFRGTTKIK  
MDGSRKKEDPINCIIISVTYPNKVGDWEWEGKEREVNFNLERVDDYERDIHVSILCMLHAQ  
LDNFEQALGENANSFYFKKGQRMFLPKKSKLFNRPTVEDSDMFSIIFPPASDQDFADDI  
YYRIIVTCS  
(SEQ ID NO: 35)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O13706 (O13706) HYPOTHETICAL 11.8 KD PROTEIN C13F5.07C IN C	36	0.47
Q20960 (Q20960) COSMID F58A6	35	0.61
CAB59514 (CAB59514) HEAT SHOCK PROTEIN 70	34	1.8
Q45851 (Q45851) NEUROTOXIN TYPE F	34	1.8
O17208 (O17208) C01B12.2 PROTEIN	33	3.1
BAA83026 (BAA83026) KIAA1074 PROTEIN	32	4.1

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has sequence identity to nucleotides 1 to 954 of CT1044: this corresponds to nucleotides 143768 to 144721 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=330  
5'stop=350  
3'start=406  
3'stop=427  
5'primer=GCCAGTGAACACGGTTATCGA (residues 330 to 350 of SEQ ID NO: 34)  
Tm5=59.57  
3'primer=ATCGTTTCCAGTCATCTCCGTC (residues 406 to 427 of SEQ ID NO: 34)  
Tm3=59.04  
probel=CTCGAATACGCTGAAGAT (residues 382 to 399 of SEQ ID NO: 34)  
probelstart=382  
probelstop=399  
direction1=Forward  
Tm1=69.00  
score1=1.99  
length=98

CT1045  
Nucleotide  
Genomic coordinates:  
Start: 150687  
Stop: 154344 (SEQ ID NO: 36)

## Amino Acid

METTMDNVVQNNDVTKPTPDVATVTTATEKRQSCKEKKDQLKAECPOVLRAKLSNTLKA  
NFGKSMSAIFAQHLVDMTNAKHFKDPKTKKILELDGSSSSDSEEEETSSSSKRKRGS  
RSASSKKEKCPNTIKNWLNDACQVFRQFADIIINLPSFDDLREVKDEQTELKTIYDLYR  
QDMEKVVEEVLGRQDLFDHKSEIAKGLARFDTHVSLPSDRSAVLDSISKELEKNSKGP  
NSNIFDTLNTLKEEIKELLCHHVKYLLQNLTPEDANFVFNSSVKYVKKSYQYYIQTSEME  
SDEFKSLLTGVNIKILEKIISSDNNVATPYKHITNPRNIISLQKVRETKPVSKDYPPFRV  
DTARDIVLLPETGGISDLPIKPVTLLQLVSYINALFSLERRNVFTDGFFNAACVLISQCL  
TNANLLSNDFFPKPIELAAVTRHNLSSMKMLQEGSSSEKSKKKEKKKDKKGGGGGDDDS  
DSETDSSSSSSSSSSSSSSSEDEEEEEKGEAVEKGGKTKRKTKKKPSKDDDLDTISKLI  
LKTGGYFHDTSSELGNKIRNLIDKDDFAGVAQYAVTITEMQSTPMNQRLVSSLLDLIMRLK  
EQVKYSVDTESTSSTAKSNNALDSAKLTSQQVVTMMVDSGAELARLAFAFFVVDNTVFN  
RHEAFILTSKLLPSNENRGLKTVVESFFKNLTISNKVSTSNEEEMSVMPFEDEQQQQQCP  
QHEQQPDLKRVVGEVFLEMGKSI VNSFPSNKS VQLTADAFKQNYSPMGRRLNLA  
ISIGSNISPNILFSLNPESVGNNTVTGLRLTNLLKNISQSAQANNIKNANTLVNNTMDQ  
QNSAAMSILLFPPTSKESTIFPGNDPSSIKLQDMTMSNLARGFYSAEGCIGVVRSEF  
DEGGVKAYTLLVDSNTMDMAVNFAAQSLEKSMSEALTNNANMNPSNVLEGGSFVDGALSY  
MFEKNGSDCEPTPLAKYTMKDVSNRYLKKFNNDKNTQDLYKNRAERALVEQVTNKPTS  
HSQLANAMGVAVIGAASIKLMEAEAAESEMRAANYQATSKSTNAINITNTIGMIRNTTHL  
CTTIAVSAADMSKLANHFMFSVLNTANNSSHRRGRSSMLLQQQPTHSFAFLEQTRGR  
GGGVLGSGTEQTKDHVERMKRDWILNMISPEDKNTTTTTPSNAGRTLGYGSNITGINTIK  
QDDKSMMDKLEMSSEFRT  
(SEQ ID NO: 37)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q9XYZ6 (Q9XYZ6) HYPOTHETICAL 75.5 KD PROTEIN	59	2e-07
O55035 (O55035) PEPTIDYLPROLYL ISOMERASE MATRIN CYP (EC 5.2	59	2e-07
VTA2_XENLA (P18709) VITELLOGENIN A2 PRECURSOR (VTG A2) [CON	57	5e-07
Q07034 (Q07034) RNA BINDING PROTEIN	57	9e-07
NAB3_YEAST (P38996) NUCLEAR POLYADENYLATED RNA-BINDING PROT	57	9e-07
O95367 (O95367) CBF1 INTERACTING COREPRESSOR CIR	57	9e-07



## Comments:

## TaqMan Primer/Probe Sets:

5'start=1624

5'stop=1646

3'start=1702

3'stop=1720

5'primer=AAAACAGGAGGTTACTTCCACGA (residues 1624 to 1646 of SEQ ID NO: 36)

Tm5=57.88

3'primer=CTGCATATTGGGCTACGCC (residues 1702 to 1720 of SEQ ID NO: 36)

Tm3=57.81

probel=CACGAGTGAACTCGGCAA (residues 1647 to 1664 of SEQ ID NO: 36)

probelstart=1647

probelstop=1664

direction1=Forward

Tm1=69.00

score1=1.99

length=97

CT1046  
Nucleotide  
Genomic coordinates:  
Start: 154556  
Stop: 156932 (SEQ ID NO: 38)

## Amino Acid

MSLVENNTQEEMILETTVEGVVEGAEVAPRGVVKRPLPSSSSSSSSASDSEDEGGEQPQTK  
PPKKKRININSKYWKIETIEPASPEMLSAVNDIDNVSKTIPLIDNSFGVQFKKSVSEEQI  
KTLTETETIAVEYGTITNVKYSTFNQLERTGEPLKKKRSNNGNNYRYWQIRIEAAAAENV  
TQAVLDAIVEGNDTVIKAILLPEGEGIGLQFNKSVSSQAKNIVQAADIEFGQVAHMKCN  
LFHKMEKADESSNSSGESPKVKKVRRNKSQPTNSYYTFTMIGDSLQERIDNAIKVIEMSP  
VKRPFNSNSAAAAEEDTTTTTTSTGVVNPRGIKDIHFDDSSISKGCFTVRNIVAANGEVP  
QEEFVSELYTNLLKVEEKVDHPTFKKLIHDRTMNRHIKAWYCICPYTTGGVPPAADKVS  
AKGIATYRIYEDRTGVFQFDGAHTSTTPAQAAEATGAIHKSMLFQSPGTDIQKFLDAKKA  
EGLEPISSGEIVYRSKWSPNDSRATRCFKFYSSSDEKMNIADVLSIVHTDGLFSSVHFRK  
DTMEYGVAKSKSKIIPKTIKIKKGGDTFHSEEDIEVPVKFTAITSEELNRECNTKGMNSL  
RAHKKRKSNSSTTTSTTSTANTPKKTKKSASASDPFAKLTLDYVDSTSFVFNIS  
KEMVQRILAQERVKTLKAVKNEEKMEIVEGEEAQETYRGIVKIKTNAKAYNLANKTSCVL  
FPADKVCLKHTLEDLGDVLDVDFVREDNVNKTVASTTTTSSSENKASGGDDEETPMEFETD  
GEKLLHELLNE  
(SEQ ID NO: 39)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q45759 (Q45759) CRYSTAL PROTEIN	41	0.023
O97003 (O97003) L1156.7 PROTEIN	41	0.030
O02061 (O02061) B0041.7 PROTEIN	41	0.039
AAD55361 (AAD55361) XNP-1	41	0.039
KI67_HUMAN (P46013) ANTIGEN KI-67	40	0.051
P79065 (P79065) NOC1 PROTEIN	40	0.066

## Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 1 to 935 of CT1046: this  
corresponds to nucleotides 155963 to 156897 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=1338  
5'stop=1355  
3'start=1419  
3'stop=1440  
5'primer=TACCCCAGCACAGGCAGC (residues 1338 to 1355 of SEQ ID NO: 38)  
Tm5=59.36  
3'primer=TGCCTTCTTAGCATCGAGGAAC (residues 1419 to 1440 of SEQ ID NO: 38)  
Tm3=59.46  
probel=GAGGCAACTGGTGCTATT (residues 1357 to 1374 of SEQ ID NO: 38)  
probelstart=1357  
probelstop=1374  
direction1=Reverse  
Tm1=69.01  
score1=1.98  
length=103

CT1047  
 Nucleotide  
 Genomic coordinates:  
 Start: 159378  
 Stop: 161256 (SEQ ID NO: 40)

Amino Acid  
 MSCSSSSSSSSSEENEVEGVEGGGGRIGPTEAKKKILRKRKRSSVKSTSSSSSSSSSSSD  
 DSDSDREEKEGRKLYVDIADTRKPPKVRKLDTPSQTLENDLYMSSSSSSSSSSSDSSSSSS  
 GEEESDDDDDDDDPDNVHVLGCKKEKSPQDIEAEKEKEEEYEEEFKRMALPSRINTSV  
 DCVIPDRILTLFSTLLKKNSFQFSQPVSFRLRLVMKQVNEAMNSAFSSMLSSSGMRLVEDS  
 LGDTSKISSFITPQTDTSNSSSSSTFVNNCTDEDIKKRNIAMGRVAELLSNIAASSNEEN  
 NFRPVVSLMRGPTCGGSNASNKKLNSNRQTIPQVLNKVIFFREIHSVIALYLSSVCVQRA  
 MNNDNTNSSGYAEGMVTIKILNIIGKIPYNEMSREKFISVGRDALYLYQNVITDMTGPKHN  
 KRLRIPQQQADFCYIIAMLVNDVPITSDLLLTGKATNLVQFASAMVDPAYRLAVHKMASV  
 FNSSYSVYKVLDLDDHKMLLRANLILSILSARNKCLSERKPRTLQSVYLFNLHLLRNKLR  
 SSGLTSEESSLGTAVKLVSQLMYEGVTRQTIEDGCSMISGNFEDEDGVTLKCLGADV  
 VKTVGLSALLSDRLRKNIRNVPFY  
 (SEQ ID NO: 41)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
SR40_YEAST (P32583) SUPPRESSOR PROTEIN SRP40	74	2e-12
NSR1_YEAST (P27476) NUCLEAR LOCALIZATION SEQUENCE BINDING P	63	6e-09
VIT2_CHICK (P02845) VITELLOGENIN II PRECURSOR (MAJOR VITELL	61	3e-08
O95815 (O95815) DENTIN PHOSPHORYN (FRAGMENT)	60	4e-08
VIT_ICHUN (Q91062) VITELLOGENIN PRECURSOR (VTG) [CONTAINS:	59	1e-07
GAR2_SCHPO (P41891) GAR2 PROTEIN	58	2e-07

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 2 to 888 of CT1047: this corresponds to nucleotides 160277 to 161163 of the genomic reference sequence.

CT1048  
 Nucleotide  
 Genomic coordinates:  
 Start: 161717  
 Stop: 165020 (SEQ ID NO: 42)

Amino Acid  
 MFGSSANNFNGDKKSSSSSSAAASSDDQQLGPLGLSTADFKKVAAILANRTESLYLLPDS  
 PNEKNNVINPNQISIVPFLGSSKAAESGSANKNENQAENSSKGGSDGKKSSQONKFNLLN  
 KVEAEEMAFKRVAELIADTPPSKDNPLRDDPDAIPSRNPWVKLTQKNLEYLFWAEAVTIEV  
 SNDRSIRSGRYLQASEVGENPFLMTISVDIRILQRMALNVVWFFNRFFRMVSGLGVENRA  
 NSTYVATSDAIAQIWVEMLLKNFISGENVPQALKYLKEHYEHVYNKISKGRQPSYFVVE  
 FERVDNTIGFVNSDTEHNGSSYMEYRCFDTIRKNASSGPGSGGKSGVLSSGTFIDNEMG  
 NNNSSAAAASAPAVSAGVSPSLSPFSSDGDGDDDDDCSGDDVWGKKMIFNTSGDGSSESSG  
 QNGGGASTYKRFRCGENTASLSQKENVRLMAMPKGNEDKQLLKNIIINFLNSALNSVENHV  
 MCTDENIFDEDAEHYTSNKELYKAIVCSNPANVYRVMVELFVNILPRLRNPIVSDIET  
 VQNLPSNNGSVRTKKMVEHGCTDMRYDIPYAKGKIRLSAKRACERKCLKDVRCFDKSR  
 EANLTPSQKAGREVVEEPFPRNHNSHRVNAHDFTFYDKYRARMNKLKDSKKKVKKIDTFT  
 TTDDFLLQDRNAFDLLRKCFLSASLHHIFCPDVLVHVRGDSFNINFANNKLECYNERNGI  
 EEVTSSTQTVNAKEALEDITKIKMKRGDDIIDVVKSKGLSLREFSKKVSIVRRFNEITNQ  
 LCNNCNVNSSNGDVFHVFTSVCVYIHNIIPVLEDISIFAEELGEELTKLVKECRDVAGED  
 KTYDDIIRNYEITVKYFKLFNALVKFCHRNYNVAVTSAINRRGYMCMVSNLVGYCKLSD  
 NAIQYHESLCSLHSSISYADYYTSRNNNSEDGGGNSSEKSNADVAKTMASFYDQFDKSE  
 DSKKNKNKTSNEILIKMFQMDRVLGMDDDDDDEDSDSSSENEEEEEEEIVKKPAKKRK  
 VEDVDSNKKTLPEPAVKVKQEEDVMEEVKEAAEEEEKKEEQEAKKEEDATEYDDDTTEE  
 DEKAVASDEDEDEDSKAIF  
 (SEQ ID NO: 43)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
BAA83091 (BAA83091) HHNBV-XIA	427	e-118
Q9YTL7 (Q9YTL7) ORF 48	88	3e-16
Q18401 (Q18401) COSMID C33G8	77	7e-13
O35788 (O35788) CYCLIC NUCLEOTIDE-GATED CHANNEL BETA SUBUNI	74	5e-12
O96127 (O96127) PREDICTED SECRETED PROTEIN	73	8e-12
O96229 (O96229) HYPOTHETICAL 78.6 KD PROTEIN	73	1e-11

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 109 to 282 of CT1048: this corresponds to nucleotides 164907 to 165080 of the genomic reference sequence.



CT1050  
 Nucleotide  
 Genomic coordinates:  
 Start: 170831  
 Stop: 171461 (SEQ ID NO: 46)

Amino Acid  
 MASSSSSPVALSSVASSVMMERDEENTLSLRNRNVNKPTPVSAAWVPVDEEDEDREEMRR  
 LEDFSSDEEDDDNKSCHCDHSDDDDDDEEDPSCFKGFSAGLCSFVRGFFGFLRKSLTKKQ  
 VFLLTSAAVAAIFKTRDVAKTEEGAATMEENSTDVITGGDGDGSGIAADVSLASEGEGEN  
 GSLLESIATTLIKTTIENLVDGGEETTEL  
 (SEQ ID NO: 47)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O77320 (O77320) PFC0335C PROTEIN	45	4e-04
CAQC_RAT (P51868) CALSEQUESTIN, CARDIAC MUSCLE ISOFORM PRE	43	0.001
O49209 (O49209) PUTATIVE HISTONE DEACETYLASE	43	0.001
Q9YPA9 (Q9YPA9) HYPOTHETICAL 45.2 KD PROTEIN	43	0.002
Q98148 (Q98148) ORF73 HOMOLOG	43	0.002
O40947 (O40947) ORF 73	43	0.002

#### Comments:

EST confirmation of the predicted transcript:  
 An isolated EST has equence identity to nucleotides 1 to 637 of CT1050: this  
 corresponds to nucleotides 170864 to 171500 of the genomic reference sequence.

CT1052  
Nucleotide  
Genomic coordinates:  
Start: 173177  
Stop: 175862 (SEQ ID NO: 48)

Amino Acid  
MTRHGVLPVK GRSRHVILGN VDYTFCTTDN NCVSLDIDFK DNITDQNIQL  
LNKKLGKKTAK KIKKEDAPE TKENSDEDIY ATKEFEQTIK GLQTKKGATE  
ENAIATAAAAA ATAAAVEKAM LSESEKSMV INRARMVLSK RDTSQKQFTA  
LKNRESFFSV LIFETGSVIV VGLQDPSLTK LCVIKATTDI ADILQKNISV  
ANVSIVNTVS TFNRFHNLFI RLKGFERNK ISYSPETPF PGMFFKLRVP  
AKPLLPGETI GEYYTKVAMM RDSKDPNFKM SDWLRIKTAL TFKVKGITVL  
GEGESGCGDV SVVSKLLFGL FHYFMDNNIK MSPKEAQRVR EKYGIPHLEW  
YLYIDMLLHS YPYVKPSAEQ VKRAMVDQOH ISEVDRITYG TKNSMDAAMS  
ANLVPSKEES ISFIKKIRSQ QLFGLHCKPS KETTRRAIDT LSFDPINQDR  
WWNKNDQYYG KERCDPFSSVA RLVSSENIN SMMNSRISCQ GKWWLDENEY  
KDKLDHIVDL CTEEIVEECE SKGFIAFPFL RKHQKEKIPT PYVLLARACN  
QKNGNKMSIN NNSNYLSGSS RAKRNAKLQE KHRVTLARLN TMMASYRFLN  
NYISTDIAPD FAKLFGNDVY SLHLMTNLP KSRGHALTYN ERALSSNEST  
YKTPGNAYFS TLFEKSIINN QETANKGNRR KRKFSRIGQE KSSFLCNACG  
VNLNKGSDIE IKGICTSCDQ NSTSYIENAL SDINRDKKIK RFKAAATHPP  
VKQELVDSLS SSSSPSSSSS QTSNKNRRT PSDFIDYVYK FTDETTGAPK  
VGLVFKMCDI LASLASRRGM EDRPTANYRT SLHSATQNKI NLNKLVSIAI  
KETGATETEA QIFNKIIGSE KGLSILCQLV ERRNKDNVVF D  
(SEQ ID NO: 49)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
gi 6647869 sp O29874 TF2D_ARCFU TATA-BOX BINDING PROTEIN (TATA-...	44	0.001
gi 6647868 sp O27664 TF2D_METTH TATA-BOX BINDING PROTEIN (TATA-...	43	0.002
gi 3122925 sp Q12731 TF2D_EMENI TRANSCRIPTION INITIATION FACTOR...	43	0.003
gi 135643 sp P13393 TF2D_YEAST TRANSCRIPTION INITIATION FACTOR ...	41	0.010
gi 135627 sp P28148 TF22_ARATH TRANSCRIPTION INITIATION FACTOR ...	39	0.038
gi 135626 sp P28147 TF21_ARATH TRANSCRIPTION INITIATION FACTOR ...	39	0.050
gi 3334374 sp O43133 TF2D_CANAL TRANSCRIPTION INITIATION FACTOR...	38	0.065
gi 135639 sp P17871 TF2D_SCHPO TRANSCRIPTION INITIATION FACTOR ...	38	0.11
gi 2833518 sp Q57930 TF2D_METJA TATA-BOX BINDING PROTEIN (TATA-...	38	0.11
gi 417896 sp P32085 TF2D_CAEEL TRANSCRIPTION INITIATION FACTOR ...	37	0.15
gi 417882 sp Q02879 TF22_WHEAT TRANSCRIPTION INITIATION FACTOR ...	37	0.19
gi 121568 sp P22010 GR78_KLULA 78 KD GLUCOSE-REGULATED PROTEIN ...	36	0.25
gi 3122941 sp P93348 TF2D_TOBAC TRANSCRIPTION INITIATION FACTOR...	36	0.25
gi 3915894 sp P52653 TF2D_ENTHI TRANSCRIPTION INITIATION FACTOR...	36	0.33
gi 1729908 sp P50159 TF22_MAIZE TRANSCRIPTION INITIATION FACTOR...	36	0.33
gi 1351224 sp P48511 TF2D_MESCR TRANSCRIPTION INITIATION FACTOR...	36	0.33
gi 1729907 sp P50158 TF21_MAIZE TRANSCRIPTION INITIATION FACTOR...	36	0.33
gi 135640 sp P26357 TF2D_SOLTU TRANSCRIPTION INITIATION FACTOR ...	36	0.43
gi 1709903 sp P54637 PTP3_DICDI PROTEIN-TYROSINE PHOSPHATASE 3 ...	36	0.43
gi 2833459 sp Q55031 TF2D_SULSH TATA-BOX BINDING PROTEIN (TATA-...	36	0.43
gi 586175 sp P32623 UTR2_YEAST UTR2 PROTEIN (UNKNOWN TRANSCRIPT...	35	0.56
gi 2833446 sp Q52366 TF2D_PYRKO TATA-BOX BINDING PROTEIN (TATA-...	35	0.56
gi 3915739 sp P18428 LBP_HUMAN LIPOPOLYSACCHARIDE-BINDING PROTE...	35	0.56
gi 135634 sp P26354 TF2D_ACACA TRANSCRIPTION INITIATION FACTOR ...	35	0.74
gi 135636 sp P20227 TF2D_DROME TRANSCRIPTION INITIATION FACTOR ...	35	0.74
gi 417102 sp P32103 H1_EUPEU HISTONE H1, MACRONUCLEAR	35	0.74
gi 2833477 sp Q57050 TF2D_PYRFU TATA-BOX BINDING PROTEIN (TATA-...	35	0.74
gi 3041729 sp Q03410 SCP1_RAT SYNAPTONEMAL COMPLEX PROTEIN 1 (S...	35	0.74

WO 01/38351

69/201

PCT/US00/28888

gi 3122946 sp Q42808 TF2D_SOYBN TRANSCRIPTION INITIATION FACTOR...	35	0.74
gi 126745 sp P14873 MAPB_MOUSE MICROTUBULE-ASSOCIATED PROTEIN 1...	35	0.74
gi 4033393 sp P78695 GR78_NEUCR 78 KDA GLUCOSE-REGULATED PROTEI...	35	0.74
gi 126511 sp P12744 LUXB_PHOPO ALKANAL MONOOXYGENASE BETA CHAIN...	34	0.97

Comments:

TaqMan Primer/Probe Sets:

5'start=1268

5'stop=1291

3'start=1327

3'stop=1350

5'primer=TTGGACATTTGTGTAAACCTTCAA(residues 1268 to 1291 of SEQ ID NO: 48)

Tm5=57.29

3'primer=CCTGTCTTGTTTATAGGATCGAA (residues 1327 to 1350 of SEQ ID NO: 48)

38)Tm3=57.45

primerScore=0.74

allele1=

probe1=AACTACTCGACGTGCTAT (residues 1296 to 1313 of SEQ ID NO: 48)

probe1start=1296

probe1Stop=1313

direction1=Reverse

Tm1=69.02

score1=1.97

length=83



CT1053

Nucleotide

Genomic coordinates:

Start: 175839

Stop: 177107 (SEQ ID NO: 50)

Amino Acid

MSSTDLSKNAFHDWVVSKTDCVFDVHCETDRDCGAACENTYSVDGKEVTKFSCNQSGR  
 CARSVYSASSLERAANDLGHIIGI IKKNPKLEEELPESFLWFINHNGGDLFVNKRAAYYD  
 TMHLSIGKLDNVDTLAQGLDKRMASSLREHLLRKLD SILLQIDKVYKAKKWILDITQE  
 AGTEEDNKEEEDAKKEDQSLSVSEIVDVLGTGTHDPMPLRARGFIQKKIYPLSRNELRELA  
 LKELFPEETTSPOVLSRQHDVSTREDLCNESMNAGRAESIFSDPDSGEYVATCACLYSEY  
 LTGPACKHKTYRYVIDYDKWKRTGRPEFLTDPVLHFKKAEAVCKSTNPNLRAIYSPDNKG  
 FLCAPVAELVKTALTFRGSHEPSLIVERDINQAENLPSNSFGVNWPYVNLLNRIQDQYT  
 (SEQ ID NO: 51)

Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q26648 (Q26648) TEKTIN B1	37	0.30
RRPP_VSVSJ (P03520) RNA POLYMERASE ALPHA SUBUNIT (EC 2.7.7.	36	0.52
RBB1_HUMAN (P29374) RETINOBLASTOMA BINDING PROTEIN 1 (RBBP-	36	0.68
CENE_HUMAN (Q02224) CENTROMERIC PROTEIN E (CENP-E PROTEIN)	35	0.89
RRPP_VSVIM (P04880) RNA POLYMERASE ALPHA SUBUNIT (EC 2.7.7.	35	1.2
Q89487 (Q89487) PHOSPHOPROTEIN	35	1.2

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 1109 of CT1053: this corresponds to nucleotides 175933 to 177041 of the genomic reference sequence.

CT1054  
Nucleotide  
Genomic coordinates:  
Start: 177123  
Stop: 178524 (SEQ ID NO: 52)

Amino Acid  
MSASLILDEYLKKTASAVLDVADSFEEKIKGEIQSPPEAAALSVALYGAPPKPSASAVASII  
TGERTSLNDKYLSDNVLLKMSVARVQGQENNRKRADQAADEIRTIMEDITGSLSGAYRQYSP  
LEEENKVHIGIMNNKTPSIVCGYYTMDTSISSEPLSLTDFQNPTVIANVTKRMESIFSKVD  
SARSTRFADFVNGVANMMDIKSSIDWANMVENVIKLPDSTPNPCSVDTIVSRDASVVKTAV  
NDIYASVGKSYCRPATQLTFMSEIEKLRKAAVVCFEALMSDTRERAFVEFLFYVSFKEDAS  
NTNSKLFVQNLSSMSGNPRQPIKLVRRSAEETLFGLCFMFKVMPPEFMNCIFNFPTIPHS  
TOYHGLYGTCLTPLLRKYGSSFEKSWAHFEEILSERANAVKKFGVNDTRIDCLDAVANLTG  
PVYVLILDVRLTSAQRSCSTKFLREIKENYLLWNRVFSX  
(SEQ ID NO: 53)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
gi 126194 sp P17279 LEU2_RHIRA 3-ISOPROPYLMALATE DEHYDRATASE (I...	33	1.4
gi 6685598 sp O95613 KEND_HUMAN KENDRIN (KIAA0402)	33	1.4
gi 2497227 sp Q04893 YM96_YEAST HYPOTHETICAL 113.1 KD PROTEIN I...	33	1.4
gi 134393 sp P17065 SEC2_YEAST PROTEIN TRANSPORT PROTEIN SEC2	32	2.4
gi 547908 sp Q02455 MLP1_YEAST MYOSIN-LIKE PROTEIN MLP1	32	2.4
gi 2496893 sp Q09462 YQ52_CAEEL HYPOTHETICAL 30.9 KD PROTEIN C1...	31	4.2

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has enquence identity to nucleotides 1 to 899 of CT1054: this corresponds to nucleotides 177561 to 178459 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=678  
5'stop=701  
3'start=754  
3'stop=777  
5'primer=CACCTAACCCTTGTTTCAGTTGACA (residues 678 to 701 of SEQ ID NO: 52)  
Tm5=59.08  
3'primer=CAATAAGATTTTCCAACAGAAGCG (residues 754 to 777 of SEQ ID NO: 52)  
Tm3=58.08  
probel=TATTGTGTCCAGAGACGC (residues 703 to 720 of SEQ ID NO: 52)  
probelstart=703  
probelstop=720  
direction1=Forward  
Tm1=68.92  
score1=1.92  
length=100

CT1055  
 Nucleotide  
 Genomic coordinates:  
 Start: 178529  
 Stop: 179348 (SEQ ID NO: 54)

Amino Acid  
 MAQTSKMGNTNKRCEEEVEEERQQPFTKSKSEPPSFEDKSSSTSSKKKSKSNKHTKTKE  
 EQLLEFVKDLERSDPTVPDEKVKQVEEEKSPEAIAEIFSMFGIAQDSKFKSLLPIERIKS  
 ITTKIVIDAINQPVKMLVDHLYHFKEMQNVVEKYKDDSDKLSVILKSKKSPKEFDLSF  
 SDYVDRLNRILGVIKRVAGAIESKELLQSNMIMNSVLGTVVSNIPYNMKINICVFLTN  
 FICTFANDDLTYTFRDDEKFMVMSQVTRYISKD  
 (SEQ ID NO: 55)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
001761 (O01761) C. ELEGANS UNC-89 (GB:U33058) (NID:G1160355	45	8e-04
Q17362 (Q17362) UNC-89	45	8e-04
Q17595 (Q17595) SIMILARITY TO MYOSIN HEAVY CHAIN	43	0.003
O31329 (O31329) ERPM	42	0.004
FKB3_YEAST (P38911) FK506-BINDING NUCLEAR PROTEIN (PEPTIDYL	41	0.007
AAD55361 (AAD55361) XNP-1	41	0.009

#### Comments:

EST confirmation of the predicted transcript:  
 An isolated EST has equence identity to nucleotides 1 to 661 of CT1055: this corresponds to nucleotides 178612 to 179272 of the genomic reference sequence. Hit to public sequence gi|6856160|gb|AF173992.1 to CT nucleotides 647 to 819 of nucleotides 910 to 738 of the public sequence with a 100% homology, a score of 343 and an Evalue of 6e-97.

#### TaqMan Primer/Probe Sets:

5'start=269  
 5'stop=289  
 3'start=340  
 3'stop=360  
 5'primer=CCCCTGAAGCTATTGCTGAAA (residues 269 to 289 of SEQ ID NO: 54)  
 Tm5=58.07  
 3'primer=GCTCTTTATGCGTTCAATGGG (residues 340 to 360 of SEQ ID NO: 54)  
 Tm3=58.32  
 probel=AAGTTCAAGAGCCTTCTT (residues 322 to 339 of SEQ ID NO: 54)  
 probelstart=322  
 probelStop=339  
 direction1=Forward  
 Tm1=69.00  
 score1=1.88  
 length=92

CT1056  
Nucleotide  
Genomic coordinates:  
Start: 185432  
Stop: 186830 (SEQ ID NO: 56)

Amino Acid  
MAVG DYLSMSSVGEATLVGFMI LNFINFVTILSLIIYAVTDVYRRCKRPSTNGYSGCTTN  
VVSSTLQEANLVTEKDKPVQFVRGLVPRKMMKEYRSDLSPKNVGEYILPSEKETDKLKS  
DYKKGKKVGLLTALSNGHDSNKRIIGPRDLISRDDVKDSYVFKRLSKDPLVYSSATSK  
YVRKFSPFRAKKFMTSTQLGSKLVYPPIRYGTAFVLPTGYVINKAYGMDNEDLHTWNPP  
SSSVLV PDSNNDRLTVECAKTDPTHRIGIYGFGGSDNRRAKEEGYVEMLLCNCNDNHKDL  
LKAPLITEYSTNPTEIQVDVAAKRVLFPA PGSEPVKSSQVTSAAHQLDGATGEHDISHEP  
VKLSDTG DYAVGSPIVFKPVYGTSLVNL PETGSPALNCPCTDKADGIYQVNQKGGILYR  
DMVGYLNANPVEAASLSSSDSSSWLT TGNKISSVTCEGEKIKKIV  
(SEQ ID NO: 57)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q99175 (Q99175) HYPOTHETICAL 58.5 KD PROTEIN	34	2.2
CAB49723 (CAB49723) HYPOTHETICAL 52.6 KD PROTEIN	34	2.9
Q9Y6X0 (Q9Y6X0) SET-BINDING PROTEIN (SEB)	34	2.9
Q9ZXH3 (Q9ZXH3) INT44	33	3.8
Q55105 (Q55105) MULTIPLE LIGAND-BINDING PROTEIN 1 PRECURSOR	33	5.0
Q9WXH8 (Q9WXH8) PYRUVATE ORTHOPHOSPHATE DIKINASE	33	5.0

#### Comments:

TaqMan Primer/Probe Sets:  
5'start=663  
5'stop=684  
3'start=705  
3'stop=725  
5'primer=CGTGATCAACAAAGCATACGGA (residues 663 to 684 of SEQ ID NO: 56)  
Tm5=59.59  
3'primer=GAAGAGGGTGGGTTC CAAGTG (residues 705 to 725 of SEQ ID NO: 56)  
Tm3=59.41  
probel=GGATAATGAGGATCTACA (residues 687 to 704 of SEQ ID NO: 56)  
probelstart=687  
probelstop=704  
direction1=Reverse  
Tm1=68.99  
score1=1.88  
length=63

CT1057  
 Nucleotide  
 Genomic coordinates:  
 Start: 190875  
 Stop: 193236 (SEQ ID NO: 58)

## Amino Acid

MEYMEEGDIAERRSEGVYILDENSACVVNVKSIRNRLGAMDAEEAQYAQDISAQLVTHI  
 IRLAHCSESNIKIDTIASIAGLFINNIFDNNSTKNKLKTYNQFKAESQNKSSVLNIFGSL  
 DPLSMLSSFMGSDPAKSGGENLDKSLGVLFVQLQYNPCKIDDIVLLEMCPKCAACTGL  
 KEAIRQEQPMEAMLLFFKCINHNRFNFGSDIKSAYASETCMRYSQDERAVVVPLRSILLG  
 CLDRDDPAHTLSSFGDTIEYADSDNAWVSSLFAAVSRMPMVDRAVIAHFYVYTMLSRHRR  
 VSGDSFKQFVYTVFVRMIYSAIEILFCDTENSSVECDGKHFLSYVNAMVNVSVLGSTFNV  
 LKAYRSWVVDQASVAPVLDIISGGWKKNYPSPDHIKRVAYDISQVINHLASPSRMVKGNN  
 KASNVTSGLDIRSVRQAKEYIPFGILENKAGYGVINIAXHNISRPAREQSNGRNFCNA  
 LHILPSIKGCEALGAQKGSADQTVNVFDNFVASHMDIAMKKQGSGLLGLTSMIDRQGL  
 TTSFSPSEAEYKKRIHDFTRYVIFSSTPINDELVNSRCILPHSNVLNSPISLRNIDPESV  
 PDTRFHFLIMMWQRPNIDEPNLSALTTSQLELLLSKNQKWDKLTTRAFFNIDRINFQMA  
 AIIKNVSGSGFLDGSKTASSSSSAPNFFQIFSGAECTAKQLQSIRKFIGESMQHVQKEWS  
 SAVNNGNRGVENYDGLNAQFSEELFELLYKLIIEEDMRPSSLIASSEFLSNYVNAMDELL  
 IRANAS  
 (SEQ ID NO: 59)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
CAB52133 (CAB52133) RCOP C7 (FRAGMENT)	34	3.0
IDHC_SOLTU (P50217) ISOCITRATE DEHYDROGENASE [NADP] (EC 1.1	34	5.1
Q23615 (Q23615) ZK822.4 PROTEIN	34	5.1
YSW1_CAEEL (Q10017) HYPOTHETICAL 63.8 KD PROTEIN T25D10.1 I	33	8.8

## Comments:

EST confirmation of the predicted transcript:  
 An isolated EST has equence identity to nucleotides 1 to 636 of CT1057: this  
 corresponds to nucleotides 192639 to 193274 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=1214  
 5'stop=1236  
 3'start=1281  
 3'stop=1303  
 5'primer=TCATCAATCATCTTGCATCACCT (residues 1214 to 1236 of SEQ ID NO: 58)  
 Tm5=58.13  
 3'primer=CAGACCTGATACATCCAGGCCA (residues 1281 to 1303 of SEQ ID NO: 58)  
 Tm3=59.10  
 probel=AGGCTAGCAACGTTACAT (residues 1262 to 1279 of SEQ ID NO: 58)  
 probelstart=1262  
 probelstop=1279  
 directionl=Forward  
 Tm1=69.01  
 score1=1.98  
 length=90

CT1058  
Nucleotide  
Genomic coordinates:  
Start: 209615  
Stop: 227849 (SEQ ID NO: 60)

## Amino Acid

MDQYPEVRDTPQTEQEQAQAQQQQAATTTAAAAAAAAPTQYSNTVSAETLSAISEDGKLE  
RSIAASCWINNLPDEKMAQRVQFHPLSSTTTYDSENVNPGSSSVFLKPRALPTGGTCLA  
PNYIAVPTLRAASEIIDSIASSTSLYQCSMFNSWNLIPIFMSNSKHSQFGDRVIKSMIRN  
CFSKQKNVENLLKELRRRKVNAAKAFSHAVQQKSAVNTALAAWNAGSAANLEKLVDFCKL  
KYSFDRKXKAGGLFSASATAQSQSGTSSSSVEHTSNDFLDLILKRHKGTSLDLDSATNTF  
DTALSRVFTEFKEQARAADDAADSDHLSASDPIFSIVRHNRRREGILDSVPNIGMLAPR  
SKYSVAEYLMADRDESADIAAKIGTKIATDFEALRGDNNKRRADTSVDDLKESLADSIEK  
TSIKNTGDISVNTIPTDTEEYEFSLHITQLFAQAFLETMGSLLSCAFGVQFPFSDEGFA  
AIERIIRKTDPTDGKVSMDPSSLSQYLLLVGNFQVSPFHVSDEPKDIVFGRQVTPNTPIL  
LSIITRSKNOKNETSTIINFRDRLLVNDTVLRDATQNVSTSTPSQRRVPTAAGEPKKPMPL  
SGCLPIIRGPQVVTRESDDMISGLVGDWYISLGVYYAMGSSAAAIAAGHQRALASAESIN  
SPMMKKFSKKGKYTEEEKRIKKAMRRNADRSARILALLGQTDQYGYVEHNSTLDSFWS  
SNAAIRAKAKEDALSRAEILAVRKQLDGKCSSRDEYSMVERYLRDSFFRSVNRSGGGYE  
MFDQGFDMGRFADFLSDNSAARNAWQYAEVMRGLSKHEKRVFNIEGLFSALNSFKFPLV  
PEQGRKKTIVGGRHRLNNLKAANKIINGITEMTLQSAIDGTGISDIIIGSVSDGWGNTAQ  
SRVKALKTLNFSNGNVVSIPIVSRVKAAGSRGGETLKCVDIPSVIIANLISDKRILD  
QLCGGGMNLAEITNFIETIAGKEHTGKESVFLSPRLSVILLRYIWFNAAVVSLTDSNIK  
MPLNTMSEGTDGDDIYRDYLAIRGMVNNYNSSLSISVKAISDRYNCGSGNTSTSNKNVTI  
KTQCELLTVLQQTANALSAFTNKGVGATPDAANMANVISPIANADVVKNTNVVSGLDR  
ITETINFFSFLSQIKTMNENIEEYLRRLGEGLDKKELDNFVYPIAAIVKRELGVS  
ALSSNLDTRPITIDLNTEQPLIVKASKGYASNRYAKLFNKTRTAAEQAQMEQYNAQMA  
ANTIPQLVNRLTIPGSITADTAINVVKAFTENGESNAETHLGMGNAINEMQPLETDFG  
NVANKRLTVNVGVSVKLIQNLTVSLILAHASKAPYVFKPLVQDFAKLLAVTAETSLV  
RSQKSFPIPPSVFSSGGLFKIDREMFDMKTDYVVEVIRQLSKNATAAIERCNDSDSA  
ARIAKSGEIYNKDVASTTAAPGTSSSALTFLANNLQNPVKVSMGALPHFDMAVVPKLHG  
ISHDQMFRLSTYYQGIHKMELNSDCKPEEWDNSLPGNRASKFFGLSSVSDNNRSFNLALD  
TLASPAEICDLVTREMVKTSNDIVHNIGSNSNTDALQKSLQVGASAVEKYDESTLSTKE  
TDVYSLVSLAKSKSPSSSSSLSEGHLSKEIDRTWNTPALGTAKTTSYSVSEDALN  
APLSAVLDFRRNVVDATKSLYEVAAVCSVMSKEEDVRSSSRKIMGVMEQESVPMQDIGID  
RIASLVSTVATPKQHRRFLQTVNDYKNYLIRKVASNPLLSSRLGGISPTSGNTDYNLKA  
YDGVVSSSSSMTFSSMSVSDRFWSGVFSQCLETGSPMFADAGHGSNMFIITAPKLYGSR  
VNTYALSSGLVLRDSISSATQERKNRIAKSIEALETFTVDVGGDTLDQLRKAQNMYN  
KLSDITSNSIYSDFGNIDCAKIMKNVTSKMTARQQSDTILSSLLHELGLVHKQPPOLA  
TQFALASHVIKAKYVTNDLNNIHEKETFSQLMAVAGVADYYNVSAAMCQRLVASDVTMF  
LGGTMLQQGLFVSFLLNNVLFVQVSDNIKMNELNDETKSLLVKLVGFCGTVSDALGSRHV  
SSIRRVQNEEDKKLDRSFVTSLSAYRDLRKKTELYRETDTINKLFHQNFMSYESSMLK  
RTSLVHDVAVSGPRPRYSTLEDVLEAPSTVHKSFMVSYPERAAASRRVKRAGLRALADNR  
MESLYGEEVLNMRSSAVSSEMMDIEYEGGGFMMMSIDDEDDIAFIDSEEESESSTDFSS  
SDEYSDSSDEYDFDDDNNGQSPYSTTSYSYDALORLNSAAKPLTAIYGCRGEGEDEEND  
LYEEEQERRRRSSKMGKILRDLHESDDDDDDYDFDEFDGERSMSETIATRRAGRIQYGP  
GFLSHSNILNRPKARAFLTRGKKFRPSAYDRFFMEDDSDLLFSESTSSSSSDSPFSS  
FSKGRKCKRRTSEDQCAFVKRVVRAFPVTRVMTINGRVSMITPVTSENTVGFYENYQAN  
KRERARLIEEYKIVKGASATLPDEYVEGRASKQVSPRELRRSLIKAAAYVARTQESNLNI  
IFDALTTTSNATLVNDPSTLLGDTLLFAKQLEAITERRNRLMKDLTEISPSLFTSFGDAS  
KDTQMADAKQIVSGGNFKSAGYLGVLRLTASCIGKNTVDRLLATKNKNHLEWMTTAA  
IVFARSFNDTTFHALEDTLKMTSALTDMYSAFTNLVGEHSORLKVKSTLLDSIFNTRMA  
HTEAVMGLVYPTAFINHEMPDQYTORREMQSLALNLRGVNCSQLPRKDIDGTAGLLTFI  
TSRKFAQYCGGERGGLSLYRMSIVDALSCPSDNRLKGAVSLEVGKQWDMGEEIFYKRSNDL  
VDFCSKNNISLENVAGPIARFVNPNGTNMADIGMTDIIISRTVKDDASMIIRLRAEEGAGAA  
GKFITASAMGNLYGGIDTVVNLTEKLYDSFVLLQDSDFNTPTMATAIINRMKSRKHKA  
LKTPFGGDIATYKNFSSSEAIIVRAKEMRNSISTIVMDISKSRGINSFSSRSGSTLAKI  
STSEFERILETSAVLSNTKANLRTIENRLAEHYNKQKQFHSIINDGLSETRAVVAVIAES  
LTPVYADDTSERGASVSELLTONTLLKFIVQNELKNIIEAKRHVTAAIEGSSQLHEKMLS  
LLVASADINRNGNLECKKLTGNSNFVPMNTNDQGGTFIKHKETGIWLTDEENNTSS  
IKDNDQRRVAKTILAIVEDNRNATIRSRLQSLCFGKYAMNDIFALDDADIKNMDKLIKLE

GEALAEKASPPSSSAISSSSSSNTTSSSSSPSSSPSSSSSSFSMDYSNNLAKTIPYMPIVF  
 QNKQSNVNSSDASSSPSSSSSSSANIDNVEHKKVALQQLQTQESNDLSNVLSVTTKHRF  
 ASHNQAATVGI FNGRQHAETVVAIPNANKANNNATVSAGQGILTRFSAPENVSSSTSMQLP  
 PSSSSSSNGDDNKVPVTVRLNQYANSILSSIENASEFKDLKEAERKIDLAIQAASTTETK  
 EMVTVSKCPANQTAITAIISQAKSLKSALELLERVIKAVEVYTPDSSIAAVSLPVNGDS  
 MVSSSSGSGSAPSSSSSSSSSSSSSNVTDYFNYAYGKLKNIDENTEEGAETVQKNMVEQD  
 AAVRIPLLVSYPFSEMMRRAIDKLNEYQYLIDAIKTKIVSDTKQASSWAIKETDKELDM  
 DKEQVISKINNQQNFSNESDKIKMAISVLDNKRNELELQNNKTRSFIEETKSRIEAGGG  
 DVANFKEIIDYENTSENDNNLFQSLKAFADNSGTVYTPDMSNGRDTKSDSKFVDMYNK  
 QILEGGIKLINEGQNTVKVDFSKALEAFPRQSGASEPVSSSVERRQRERLQAVEMFMA  
 IMMERTESLRKRLADSAQWNTVNNVEETVNSGMVNIKSERLTEIRNQAQIAESTALNSI  
 NDEIVESPLTSLGARVDQLLIKVDVVGSIQQQQQQQQQQQLPKLTATEQRKEQQYAAD  
 RVVYDPSYTCFLOPHETIKRISSVYNSKNKGPLSNTRGVPTSDADLQMLTITDLSRSLV  
 DSSSTSSKKMLYENVPSSIVPGLCQQCAMMITNVHEATHTSPHSFNFENKRSKQLTEML  
 NAATSSSDGPAVRHVDLTMLESNNGYVKDFGFTHRQKVACITPVNTLLGGTFSGNVAPNT  
 VILPTSELFNCPGVENDKFRSMVNRTTDKNVADAPKSSASIVETLARTSPNAEHLYFPFK  
 DQRRHFNSITDAIISGMSGESSQLNTTCDQNLVNIQTTGFPVFTGRKQGERRIVHTEN  
 TMEGARKDKNSGIPSCTKDRQTYIDMGTKFMVAPGSLNANKEETLRLNRLSDINNVRHY  
 GTDVHVAGANSARWIGEVVRAASSFPDGDKEAMKMLLLGSVSASISQKSASHINDPTA  
 LLSTNTSIQNLVKEAFPDPVCSSNYLGSAAESTFATQLAYRQRLFPNGDDENVTTVSNICP  
 MDLMGSTKRYNDAFNIFGSKMTSTNKKGSNCENLLKSAMSNVPAINATFGAFEEASSSV  
 RNRLSPLYEDSTKYSSNQLAVQAMTDTAVDALSAVSTVVRQNGRNTLLSLPTSITSIAT  
 SGRPSSLSYSSDMKSNLIKTISRINRDASLLSMGDSQVAAGSSFFNSFLRSSIPVTTSDQ  
 GNVAAAEIVLGTILDKTVEINKRFEMLGGGKMVAGSPEARAIQRNTMSSILQMNENELAR  
 DLCEIENKIETRQLRDAFQDLKRSMLMTPGGVGAISSGASTNNVPLSLLMSRVDASSGLL  
 MNNNSANVMEAVDSFNTTPLLVRHMLDSGKSPVPMKEIRSMLTQPRALTARALLSESS  
 PLTEICLYNTRDTPERAVDRLLTSAYLVKQAKRFDGVDPAFPAAALTCASHMLSSMDS  
 HTKSSFMDNIKLHMTDTQCFKNIERFEKFLGRYGYDEYAMSHKQNCNCFHLHHTFTPSD  
 NEHLVSSFAFARPEVSMEEIRATPYQANKLISDKHYVMNMSKIDSRVTGSSLLKKVSEWT  
 EMRMNSNFNGTFEPSRLALSNSGMMTAGVNLDVIVKPNNARSVLGILECHROHVCTADAK  
 GTVASAMPAVFQATDGNNGNESELIQNALPRNRYIQKSTMNAQTVVFANVLEQLIADLGKV  
 IVNELAGTIAESVPESVYENTKEMIDRLGSDDLFKSNNNGGVESMDYEDSETTSNNGPVL  
 ISEAMKNAVYHTLISGKAARPENVPFASCASGPLAFDFLLSKGDTFEEKNAEQGAAAVS  
 STYSSSSNTTLRKHLARVFEAISQVTDAEFKDILNDIERNISSDYTNCPPTNQNAFAL  
 ATKREFSRIVSFLTILRKNITPALVDPKGALHEKVAIYLTLLSTKSKLENFFQYGLSNSS  
 SVDLSHLKPINCNNVKNIEDTFMYRNVHPILIMALPENFTALLQQEQMDPDTAIESRRS  
 LTTFNLHPNTASMANGARAAGGAGGNPMGLYLSSHILHESTVTTSNPVTDTTENVNYHS  
 SVTQDPVMVVPFKDSARLIVNNNTGIDVLNDKSCNYLQVSMPSSESSGLVTNTGCSSSS  
 SSSSDTFKYVRDNTPVNLPVPAVLCSDASSNLLDVFSRADIVLENMNVRFGMPEI  
 IAAVSKFKGLTKEEVIKQMVSONNNNNNGNGKKTTPVTPGTGDIVITNATFPDTRP  
 LYTAANGGTSSFKWGDINDRKMHAKEPFTFFIGNPTAAATANGVPLTSEGISLTEEKRRK  
 IAGISEGSIGTGALRAAANTRLSSDMEPVMKGWNNIVQLQQTFFKASDKLTHLLRSGGIP  
 PRSQETNAIINKMHDSFKTLEECRRVIQDEAALLVATSDLLTGYYGGDAALAMVSPVRPE  
 MTGLIGAISAPVRGISHLLKLGVSAAANAAIRKRLNLPSTNGKTLPEHGIVHKSATLLL  
 DSDSISNLYNTDLQDVVSNARDNNNLGRIMQSLGLKGNAGDLVYSARQLTDLITVPEYG  
 NNRDLTKRQAILKMLISNPEILENVADTIYLTGKNALAPVSAQEMACASLTVGGSGGGK  
 LSSDDNVQSLNRLYFRV  
 (SEQ ID NO: 61)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
SR40_YEAST (P32583) SUPPRESSOR PROTEIN SRP40	80	5e-13
O94317 (O94317) SERINE-RICH PROTEIN	80	5e-13
O95815 (O95815) DENTIN PHOSPHORYN (FRAGMENT)	80	6e-13
YM96_YEAST (Q04893) HYPOTHETICAL 113.1 KD PROTEIN IN PRE5-F	76	9e-12
AGA1_YEAST (P32323) A-AGGLUTININ ATTACHMENT SUBUNIT PRECURS	73	8e-11
Q9Y076 (Q9Y076) PROTEOPHOSPHOGLYCAN PRECURSOR (FRAGMENT)	71	3e-10

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 704 of CT1058: this corresponds to nucleotides 227240 to 227943 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=8829

5'stop=8848

3'start=8894

3'stop=8917

5'primer=CATTACAGCCTCAGCCATGG (residues 8829 to 8848 of SEQ ID NO: 60)

Tm5=57.78

3'primer=GCAGAACGAACGAGTCGTATAGTT (residues 8894 to 8917 of SEQ ID NO: 60)

Tm3=57.76

probel=GAGGTATTGATACCGTTG (residues 8861 to 8878 of SEQ ID NO: 60)

probelstart=8861

probelstop=8878

direction1=Reverse

Tm1=69.00

score1=1.99

length=89



CT1059  
 Nucleotide  
 Genomic coordinates:  
 Start: 228374  
 Stop: 230564 (SEQ ID NO: 62)

Amino Acid  
 MDKVCVISNTRERTFKVPADLLCVATEPEISTKEEDAGIEIETRVVVFSSRCVSVQELHTI  
 NPNDEGFVSVQLFKDYLKLSAQGKKPIGLYIYQIKAGEDLERRRLISGGTAYLDPATHLFYL  
 DFSLYPNYSIFNDISSRLKIIDEDTYNGVVFSNSEEKEKDALVLRVTFSTHEKAIEAAI  
 KKIMLRKVFFKDGDLDFGYLRIPKSKLDKFTPYFRSQYGIVNVEKNIPGYIWGEIMKQRV  
 RCSRWYLYNTDSEWEYKNVAEERVGPRQLVKKYGAKCENLCFRDIDLKKEAKEKRDIER  
 ETESRYVVVTLTHKHMPENMPYFGPKCSVVRLEDTRILLCFVDEISYNDEDVDEILSEN  
 RSLRNVSIRHKENVPVHTLLKKGVSIIHARFTLNGLDDALIILKRIPKTYFEDEELQAACA  
 HVNLEQYEWLCSNNRGNKVEHVKSRRVTRAVKRRRKRHWIYFDKDTLNLNYKYFDKKVT  
 ASMASKICNAKHDCLVFHRKMELEDLTESAYFKVEPSPINFALKKSCPDVKYVQKKTDGT  
 FSVIRFFRNMTKGDLIQRMDLFCRFIPDSHTITLLSRADFYACKRGESMHMCTNKHRIH  
 YKFSNAPHAAIEQITNIIISDTRGRKGIHIEYAIENVQEMYEEDGRRYEAKYTGTLTEYKR  
 NEDKTFKSLAPHLTPVKNPYNINHLYEQYGNFDEELEDKLRSGFISYDTYVTAKDNWGR  
 CATGKGACI  
 (SEQ ID NO: 63)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
YO61_CAEEL (P34600) HYPOTHETICAL 84.7 KD PROTEIN ZK1098.1 I	36	0.95
Q9ZAJ8 (Q9ZAJ8) BONT PROTEIN	36	1.2
Q01794 (Q01794) MAJOR SURFACE ANTIGEN MSG1 (FRAGMENT)	36	1.2
Q9X708 (Q9X708) BOTULINUM NEUROTOXIN TYPE B (FRAGMENT)	35	1.6
GLND_HAEIN (P43919) [PROTEIN-PII] URIDYLTRANSFERASE (EC 2	35	2.1
BXB_CLOBO (P10844) BOTULINUM NEUROTOXIN TYPE B PRECURSOR (E	35	2.1

CT1060  
 Nucleotide  
 Genomic coordinates:  
 Start: 230616  
 Stop: 231582 (SEQ ID NO: 64)

Amino Acid  
 MCTLKTYKMTTSTEISKNLSDVLSIKATGDWCSNIKTVFSPFTEGKGNLPNSLPFTRSPN  
 TTCGSREAANATEHFITVFAKDKYERKRVKRTIGFTLDNTKELTPNRYLVADVYSWQEEK  
 MVFEGFCVPPGKSGTFVRYSNEDKSFLADTGRYMKKKYDDPENKTSSGGDDDDDDDDDD  
 DDNNNVVDVYEENDPRNVFEVEKDEKYACTFSILVYRAMKKSPVCRGLLVETDGPSSHPK  
 RAPSAFNPFGGSSMLNGYGAGADALEEEDEVGVPERERITNFALKRGPATGQNFVSVKL  
 EHDGSKADLYNVTCFSKQRGV  
 (SEQ ID NO: 65)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O97300 (O97300) PFC1035W PROTEIN	51	1e-05
KEX1_YEAST (P09620) CARBOXYPEPTIDASE KEX1 PRECURSOR (EC 3.4	47	2e-04
Q83970 (Q83970) (CPV)	46	3e-04
SIS2_YEAST (P36024) SIS2 PROTEIN (HALOTOLERANCE PROTEIN HAL	46	3e-04
YB00_YEAST (P38114) PUTATIVE 126.9 KD TRANSCRIPTIONAL REGUL	45	6e-04
O77384 (O77384) PFC0760C PROTEIN	45	8e-04

#### Comments:

TaqMan Primer/Probe Sets:  
 5'start=477  
 5'stop=500  
 3'start=555  
 3'stop=578  
 5'primer=CGATGATCCAGAAAATAAGACCAG (residues 477 to 500 of SEQ ID NO: 64)  
 Tm5=57.85  
 3'primer=TCGTTTTCTTCATACACGTCAACA (residues 555 to 578 of SEQ ID NO: 64)  
 Tm3=58.34  
 probel=ATGACGATGACGACGATG (residues 515 to 532 of SEQ ID NO: 64)  
 probelstart=515  
 probelstop=532  
 direction1=Forward  
 Tm1=69.04  
 score1=1.95  
 length=102

CT1061  
Nucleotide  
Genomic coordinates:  
Start: 231602  
Stop: 232799 (SEQ ID NO: 66)

Amino Acid  
MQLILSHHLTMAGRVELVTGPMFAGKSTYLKNIYQQENGNGKHCLFVKHSLETRYGCGTG  
TIVTHAGEVIEGCTTVSSIKELISVLPEVVDVILIDEGQFFTDLVLVNRLADKGRIVIA  
ALDGTSDQQMFSPHKLKLLPYTNSIVKLASKCMICKIDTKEAPFTVRFGNDNDNNVICVGG  
AEMYAAACRDCYKKINKKKKNGKLVVLEGGDRCGKSTQAKLLLTNKNSPLYGGEYMCFFD  
RSSHTGKLINDYLTKKIELDDHAAHLLFSANRWEVCSKIKQLLDDGIHVMDRYYYSGIV  
FSLARGVDTVEWCSASDEGLPQPDVLVLLMLLDVEKCSNRDTFGVERFETNSIQERARALF  
LDLANKDEKNVWIKVDARGTIEEVQTKIINIVYNIVEE  
(SEQ ID NO: 68)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
O74528 (O74528) THYMIDYLATE KINASE	165	6e-40
KTHY_HUMAN (P23919) THYMIDYLATE KINASE (EC 2.7.4.9) (DTMP K	159	3e-38
KTHY_MOUSE (P97930) THYMIDYLATE KINASE (EC 2.7.4.9) (DTMP K	157	1e-37
KTHY_CAEL (Q22018) PROBABLE THYMIDYLATE KINASE (EC 2.7.4.9	148	8e-35
KTHY_YEAST (P00572) THYMIDYLATE KINASE (EC 2.7.4.9) (DTMP K	143	2e-33
KTHY_SCHPO (P36590) THYMIDYLATE KINASE (EC 2.7.4.9) (DTMP K	141	6e-33

CT1062

Nucleotide

Genomic coordinates:

Start: 232848

Stop: 233334 (SEQ ID NO: 68)

Amino Acid

MLPRKTLPTDENG YFVLDES LLEK VYYDNNNELIVRVGGIYMQICKSKYIFHHDDPERFF  
YSVLEDYHPIKEIVERLAEEDGVFLGPWEFLSRKQVNLQHGCYKALLSLPEDKYCNLLLP  
QQMKTNLEKMEEIQRTRLIHSRTYNT PQIELSDQLDGCVIC  
(SEQ ID NO: 69)

Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
O60678 (O60678) PROTEIN ARGININE N-METHYLTRANSFERASE 3 (FRA	31	5.4
P95966 (P95966) ORF C04027	30	9.3
Q9ZW94 (Q9ZW94) F5A8.4 PROTEIN	30	9.3
O58387 (O58387) 358AA LONG HYPOTHETICAL PROTEIN	30	9.3

TaqMan Primer/Probe Sets:

5'start=223

5'stop=245

3'start=309

3'stop=330

5'primer=GAACGACTAGCAGAAGAGGATGG (residues 223 to 245 of SEQ ID NO: 68)

Tm5=58.17

3'primer=TGGCAATGACAAAAGAGCTTTG (residues 309 to 330 of SEQ ID NO: 68)

Tm3=58.99

probel=AAGTGAACCTCCAACACG (residues 284 to 301 of SEQ ID NO: 68)

probelstart=284

probelstop=301

direction1=Forward

Tm1=68.99

score1=1.96

length=108

CT1063  
Nucleotide  
Genomic coordinates:  
Start: 236678  
Stop: 238604 (SEQ ID NO:70)

## Amino Acid

MVASTPCPGPGPVPTQELLSTNFLEAHKLVLVELLLPSYSSDVVYCDSETYTKPIPIFGNK  
SIVSTIGDYVLSNPNEDEVSYQMVSSVLEKFPLLFHCTYKTNEEDKGIPLWKKLYNKRKFK  
LLNSLLVHNNKNWTPVPAIPFDRENICDASGRSVMSEIMSTSTFQTICKNNTHYLFDM  
NMERGKQGGSFLLHFFASRKNSFTNFENEEMDSHVLSNIAKFICNEKEKLDSEFIPANGKIP  
CPDKTNDEGYIPLIETAIMEDNYPALLYLVCRYGASWANTYGDHNEKSLKAFIRNDKDC  
EIEFISDHYSFNKNVTKEEFVKEKTVECVGCLYDIEDEKRCYKLP CGHFMHTFCLSNKC  
SKANFRVCVKCFQTFDDTIFRKCPPTIQWKMGINQTNHKEMDLFNRAFDYLD FICSYNV  
KLDKSKPKHKPENKKVEELAKRTAEIEEAIAKKKEELAKRTAEIEEAIAKKKEELAKR  
TAEIEEAIAKKKEELSKYNKIEKGKRRLNNEECVKLRDISTAAINMYKEKVRINGVLLK  
DSDQELAEAKERLRKILLLEETKLDRLFRPKRVEERIFLT KDDETLAFKLAEKKTED  
IIAKNNQKGSERRDGEYTTITSHIEKLPQSTALASVCVLNE  
(SEQ ID NO: 72)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q26938 (Q26938) KINETOPLAST-ASSOCIATED PROTEIN (KAP)	74	3e-12
Q9ZIU2 (Q9ZIU2) VIRULENT STRAIN ASSOCIATED LIPOPROTEIN	73	5e-12
O50870 (O50870) HYPOTHETICAL 54.3 KD PROTEIN	73	5e-12
Q9ZU69 (Q9ZU69) PUTATIVE VICILIN STORAGE PROTEIN (GLOBULIN-	68	2e-10
O23230 (O23230) TRICHOHYALIN LIKE PROTEIN	62	1e-08
MNN4_YEAST (P36044) MNN4 PROTEIN	62	1e-08

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 362 to 1 of CT1063: this corresponds to nucleotides 238163 to 238524 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=775  
5'stop=797  
3'start=850  
3'stop=872  
5'primer=GAAGACAATTACCCTGCATTGCT (residues 775 to 797 of SEQ ID NO: 70)  
Tm5=58.33  
3'primer=GCAAACGCTTTGAGAGATTCATT (residues 850 to 872 of SEQ ID NO: 70)  
Tm3=58.56  
probel=TAGGTATGGAGCATCTTG (residues 810 to 827 of SEQ ID NO: 70)  
probelstart=810  
probelstop=827  
direction1=Forward  
Tm1=68.95  
score1=1.95  
length=98

CT1064  
Nucleotide  
Genomic coordinates:  
Start: 238658  
Stop: 239438 (SEQ ID NO: 72)

Amino Acid  
MSTCSNLLSVFGGWDWTTTFFDLVHTRQECDDKKREQDYSFFITETCKGENIGIHSYEHT  
SKIIDTGNNDSSTIEEVLNIYKAINHLENILKLNKGEKIILMDVETMILETHKILMKG  
ILPKGKNGSFSTCVRFAVNKNNERHYYPVFETEKEAFNSIQNLVDYYNEIVAHTNDQIKI  
IKACAYFMYNFLTLPFNDGNGRTARLLYSFLLKGNIGVPHFSPITHPRDQFVDTLVYFR  
EHGDGRPLLYVLLESIKNK  
(SEQ ID NO: 73)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q23544 (Q23544) ZK593.8 PROTEIN	50	2e-05
Q9ZHQ9 (Q9ZHQ9) HYPOTHETICAL 26.1 KD PROTEIN (FRAGMENT)	44	0.001
O68899 (O68899) HYPOTHETICAL 29.5 KD PROTEIN	43	0.002
Q48249 (Q48249) PLASMID PHPM180, COMPLETE SEQUENCE	43	0.002
O75406 (O75406) HUNTINGTIN INTERACTING PROTEIN HYPE (FRAGME	42	0.004
AAC96089 (AAC96089) HYPOTHETICAL 23.9 KD PROTEIN (FRAGMENT)	42	0.004

#### Comments:

EST confirmation of the predicted transcript:  
Nucleotides 1 to 677 of CT1064: this corresponds to nucleotides 238717 to 239393 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=397  
5'stop=419  
3'start=450  
3'stop=472  
5'primer=TGCGTACGCTTTGCTGTAAATAA (residues 397 to 419 of SEQ ID NO: 72)  
Tm5=58.67  
3'primer=TGAACGCTTCTTTCTCTGTTTCA (residues 450 to 472 of SEQ ID NO: 72)  
Tm3=57.92  
probel=AATGAACGGCATTACTAC (residues 424 to 441 of SEQ ID NO: 72)  
probelstart=424  
probelstop=441  
direction1=Forward  
Tm1=68.96  
score1=1.88  
length=76

CT1065  
Nucleotide  
Genomic coordinates:  
Start: 240712  
Stop: 241192 (SEQ ID NO: 74)

Amino Acid  
MEDLKSTIERVYEERVENLEQWTNTVEEEEERTVSAIDSVLEEQRALDAWEAAIKEREND  
LAVKEGISALVFNAADAKTRKELINTWIAERETSEKRRKEATSTNNQLKNQMSSLVNTTK  
TLKEKYNNKYRRSAILNMQYINNKR DYEASQFWVYTNN  
(SEQ ID NO: 75)

#### Top Blast Hits

Sequences producing significant alignments:

Score (bits)	E Value
43	0.001
43	0.001
42	0.002
41	0.005
41	0.005
41	0.005

RADI\_HUMAN (P35241) RADIXIN  
RADI\_PIG (P26044) RADIXIN (MOESIN B)  
O40947 (O40947) ORF 73  
SSP5\_STRGN (P16952) AGGLUTININ RECEPTOR PRECURSOR  
Q9Y489 (Q9Y489) CENTRIOLE ASSOCIATED PROTEIN CEP110  
Q21952 (Q21952) SIMILAR TO MYOSIN HEAVY CHAIN

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=228  
5'stop=249  
3'start=297  
3'stop=320  
5'primer=CGCCAAAACACGTAAAGAATTG (residues 228 to 249 of SEQ ID NO:74)  
Tm5=58.34  
3'primer=TGATTATTGGTAGAGGTTGCTTCC (residues 297 to 320 of SEQ ID NO:74)  
Tm3=57.45  
probel=AATACGTGGATAGCCGAA (residues 253 to 270 of SEQ ID NO:74)  
probelstart=253  
probelstop=270  
direction1=Reverse  
Tm1=68.84  
score1=1.84  
length=93

An isolated EST has sequence identity to nucleotides 452 to 1 of CT1066: this corresponds to nucleotides 241246 to 241697 of the genomic reference sequence.



CT1067  
Nucleotide  
Genomic coordinates:  
Start: 241774  
Stop: 243409 (SEQ ID NO: 78)

Amino Acid  
MFRQFCSLYLLQRRVNDNLRSTASASAAASLKGDTGTEFITGEPSPSHKMRGPSYSVLGPDP  
CEDPERVYVDIVVSILQTNNIQVTKEWELFSDKLRKLGPDWIDRSGIENNNGEGEEDGDENE  
DGGGNGGRIEDREAHRRKMMKKLSFVGREDPVAVDLPTWRENSTEFARRLTLKELCDLIV  
ECGCIKSKEELFDIFEEPWEIKEAADVRGMANRSKFTKESLIDWFFEFDTYSKCVVFFE  
AVNWYLSQASPISLVLDIYCCVFSYIRRQTFLTRAKNPSLTVASSFSPTPDTKLLAID  
ECVQHFLKSDINISQMALTERDCFFPLLTEMPRQKKVNTFLDTMKRPTLSLLPSTSSSS  
SSNNKRKRNTAAANILLPVYRSNFASTASNNKRLKTDGGENASACILIEGYANGKISPIRI  
MVRKSTIIEPVFNHLLFPVFASKDTGANILFFIKMKSFASASLLLPGLFRHPKQFLNGPC  
KWMTLAENNINDNNINSSTMWSYTLADYCPLGYTQESPQPYQTCGNFTSTTNKRLQNVQ  
PLYF  
(SEQ ID NO: 79)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
CAB38842 (CAB38842) HYPOTHETICAL 35.2 KD PROTEIN	39	0.061
CAB52581 (CAB52581) CONSERVED HYPOTHETICAL TBC DOMAIN PROTEIN	39	0.10
O84402 (O84402) RIBONUCLEASE FAMILY	37	0.31
Q59085 (Q59085) RNA POLYMERASE SIGMA-54 FACTOR	37	0.31
O80815 (O80815) T8F5.22 PROTEIN	35	1.2
PI4K_DICDI (P54677) PHOSPHATIDYLINOSITOL 4-KINASE (EC 2.7.1	35	1.6

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=717  
5'stop=740  
3'start=791  
3'stop=814  
5'primer=TGAAGCAGTCAACTGGTACTTGAA (residues 717 to 740 of SEQ ID NO: 78)  
Tm5=57.74  
3'primer=TTTGCGTCTTATGTAGGAAAAGA (residues 791 to 814 of SEQ ID NO: 78)  
Tm3=58.08  
probel=ATCTCAAGCGTCTCCAAT (residues 741 to 758 of SEQ ID NO: 78)  
probelstart=741  
probelstop=758  
direction1=Forward  
Tm1=68.99  
score1=1.99  
length=98

CT1068  
Nucleotide  
Genomic coordinates:  
Start: 243216  
Stop: 243798 (SEQ ID NO: 80)

Amino Acid  
MDDSSRKQHQHQHKLFDHVELHASRLSSGLLHPREPSTLSDMRQFYFDYKQETTKRAA  
IILLNTLLEYRTPSEWEI PFNLLLNMNNKWTLPVVKISAGIISKLPWTMKTMYEI  
VSSPNNNNNNNGDYSTCRMMVMEYPIGGLLHTPAITNKYPRSRMVTCTKGKDHQKLYDIS  
RQMFDIIEANGQL  
(SEQ ID NO: 81)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
P3K3_DICDI (P54675) PHOSPHATIDYLINOSITOL 3-KINASE 3 (EC 2.7	32	3.0
O01590 (O01590) K09H11.1 PROTEIN	31	6.7
RRPL_DUGBV (Q66431) RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.4	31	8.8

#### Comments:

TaqMan Primer/Probe Sets:  
5'start=234  
5'stop=257  
3'start=308  
3'stop=331  
5'primer=GGAAATTCGGTTTAACTCTTGCT (residues 234 to 257 of SEQ ID NO: 80)  
Tm5=57.99  
3'primer=GGAGTTTCGATATGATACCTGCAC (residues 308 to 331 of SEQ ID NO:80)  
Tm3=57.75  
probel=GAGTACACTCATTCAGG (residues 279 to 296 of SEQ ID NO:80)  
probelstart=279  
probelstop=296  
direction1=Reverse  
Tm1=68.82  
score1=1.82  
length=98

CT1069  
Nucleotide  
Genomic coordinates:  
Start: 244241  
Stop: 244856 (SEQ ID NO: 82)

Amino Acid  
MDLSFTLSVVSAILAITAVIAVFVIFRYHNTVTKTIETHTDNIETNMDENLRIPVTAEV  
GSGYFKMTDVSFSDTLGKIKIRNGKSDAQMKEDADLVITPVEGRALEVTVGQNLTFEG  
TFKVWNNTSRKINITGMQMPKINPSKAFVGSNTSSFTPVSIDEDEVGTFVCGTTFGAP  
IAATAGGNLFDMYVHVITYSGTETE  
(SEQ ID NO: 83)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O85179 (O85179) FLAGELLIN A	34	0.60
FLAA_CAMJE (P22251) FLAGELLIN A	34	0.60
BAA83944 (BAA83944) UNKNOWN	34	1.0
FLAB_CAMJE (P22252) FLAGELLIN B	33	1.8
AAC25644 (AAC25644) FLAGELLIN A	33	1.8
AAC25648 (AAC25648) FLAGELLIN A	33	1.8

#### Comments:

EST confirmation of the predicted transcript and hits to public SBV sequences:  
Nucleotides 1 to 803 of CT1069: this corresponds to nucleotides 244132 to 244934  
of the genomic reference sequence.  
Hit to public sequence gi|6856162|gb|AF173993.1: CT nucleotides 1 to 615 match  
nucleotides 323 to 937 of the public sequence with a 100% homology, a score of  
1219 and an Evalue of 0.

CT500  
Nucleotide  
Genomic coordinates:  
Start: 2425  
Stop: 1537 (SEQ ID NO: 84)

Amino Acid  
MKNSRQRSGVWRGNSCLYKSFYFSGAIIIECKKIRIIMMFLLSLILFVCFVGVVGVIFM  
SRPNKTTTTTSNKKTKKDKEKEKEDDTGAVLGRREPENRPIGRDEEGAVEDGKEEEVFE  
FEQPSVNTGSNTGGGGTGTVPGEGLPPPPPTPTPTPPPTPTPTPPPPPTRTPSPSSSLG  
EDDDDDIDIDFDNDIDIDEFLDSGEEMEEDEEEEDLDTLLSRLETGMSGEEVDFDASSAYI  
QPDPVVVKNIERSDYTLDPMESWKVLNRSEGDIREFFVDRGITNKKIKAMTEDLKEL  
(SEQ ID NO: 85)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q48373 (Q48373) CHITINASE PRECURSOR	64	2e-09
O86476 (O86476) CLUMPING FACTOR B PRECURSOR	59	4e-08
O92451 (O92451) ACMNPV ORF91	58	9e-08
Y091_NPVOP (O10341) HYPOTHETICAL 29.3 KD PROTEIN (ORF92)	58	9e-08
Q69023 (Q69023) (B95-8 ISOLATE) U2-IR2 DOMAIN ENCODING NUCL	57	2e-07
Q42421 (Q42421) CHITINASE PRECURSOR	57	2e-07

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 244 to 513 of CT500: this corresponds to nucleotides 1979 to 2248 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=419  
5'stop=440  
3'start=505  
3'stop=523  
5'primer=TGCCTGGAGAAGGTTTGTTACC (residues 419 to 440 of SEQ ID NO: 84)  
Tm5=59.11  
3'primer=GAGATGGGGTTCGTGTCGG (residues 505 to 523 of SEQ ID NO: 84)  
Tm3=59.88  
probel=CCTCCTACTCCTACTCCT (residues 448 to 465 of SEQ ID NO: 84)  
probelstart=448  
probelstop=465  
direction1=Forward  
Tm1=69.04  
score1=1.95  
length=105

CT501  
Nucleotide  
Genomic coordinates:  
Start: 7645  
Stop: 7042 (SEQ ID NO: 86)

Amino Acid  
MTMWNKIVITTKRMNWPVVGVFFILAITALAVLYIRHASKQEKYSTSHINEQFTAKQL  
PVTYLSKTGKLDKMDHLTHSDFMAYVDVHNRTKTLKHPMCTDEAGWAHFCLLASAEAYRRI  
RYGRGEFGPEKHSLAETIQSTVQDMSEPYITHIFKKNSTDVDGHGMQSVLEKNRNKIRMGD  
GKTSSETYNLSDKSISIVGV  
(SEQ ID NO: 87)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q21859 (Q21859) R09D1.3 PROTEIN	34	1.0
Q9Y3S0 (Q9Y3S0) EMDC II PROTEIN	32	4.0
AAD25099 (AAD25099) METALLOPROTEASE DISINTEGRIN CYSTEINE-RI	32	4.0
O31548 (O31548) YFJL PROTEIN (RIBOSOMAL PROTEIN L6-LIKE PRO	31	5.3
Q44602 (Q44602) PHOSPHORIBOSYL ANTHRANILATE TRANSFERASE	31	6.9
CAB52230 (CAB52230) HYPOTHETICAL 33.8 KD PROTEIN	31	6.9

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 587 to 1 of CT501: this corresponds to nucleotides 6996 to 7582 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=245  
5'stop=266  
3'start=315  
3'stop=332  
5'primer=TGGCATATGTTGATGTGCACAA (residues 245 to 266 of SEQ ID NO: 86)  
Tm5=59.44  
3'primer=AGCAGGCAAAAGTGGGCC (residues 315 to 332 of SEQ ID NO: 86)  
Tm3=59.78  
probel=ATGTGTACTGACGAGGCT (residues 292 to 309 of SEQ ID NO: 86)  
probelstart=292  
probelstop=309  
direction1=Reverse  
Tm1=68.96  
score1=1.96  
length=88

CT502  
Nucleotide  
Genomic coordinates:  
Start: 8502  
Stop: 7641 (SEQ ID NO: 88)

Amino Acid  
MSSGSINNHPSSNMDTNKMEEGEEQDFDVLELDYSKIIHDITAMLSVAAPPPNSILDASD  
GLIATASATAPAAETGNSNRMRDKDVCQLIERDIELVKSDTIEVDSIIRQLLYFGESAS  
EKNIKTNSTEKEPVYFPKEPKGEAVKLAKNTPVLDITKLDWMANICQSNKIGVENLASA  
LQSGQLIWTTFFPAAVYASLDSFYHIAIMWKLLGSFINIEALSKGSKDNLLPRDDIQVVHA  
KQEI AAMLQSRQNILGRGPSEYPPVPITAILSRITIIPLLRNFSEKL  
(SEQ ID NO: 89)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
043631 (043631) SPINDLE POLE BODY PROTEIN SPC98 HOMOLOG	38	0.063
060853 (060853) PROTEIN ENCODED BY SACCHAROMYCES CEREVISIAE	38	0.063
060852 (060852) PROTEIN ENCODED BY SACCHAROMYCES CEREVISIAE	38	0.063
AMPR_HUMAN (P15514) AMPHIREGULIN PRECURSOR (AR) (COLORECTUM	34	1.2
BYN_DROME (P55965) T-RELATED PROTEIN (TRP) (BRACHYENTERON P	33	2.1
AAB32396 (AAB32396) T-RELATED PROTEIN	33	2.1

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 1 to 753 of CT502: this  
corresponds to nucleotides 7704 to 8456 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=315  
5'stop=336  
3'start=392  
3'stop=412  
5'primer=TGACTCCATTATTCGCCAACTG (residues 315 to 336 of SEQ ID NO: 88)  
Tm5=58.70  
3'primer=TGGGGAAGTAACTGGCTCCT (residues 392 to 412 of SEQ ID NO: 88)  
Tm3=58.70  
probel=TGGAGAATCTGCATCAGA (residues 345 to 362 of SEQ ID NO: 88)  
probelstart=345  
probelstop=362  
direction1=Forward  
Tm1=68.85  
score1=1.85  
length=98

CT503  
Nucleotide  
Genomic coordinates:  
Start: 9248  
Stop: 8552 (SEQ ID NO: 90)

Amino Acid  
MDPGASAASRRALWSSTVTNTRHYQQQLNRALNKIEEEDDVEEEHGVTTTNKEMASTST  
SSSSSSSSSPTSSAIPSSDEEEEEEEYDSESDTNVDSLGEEDSDTESTSADANFL  
RSSRNSTTRNRLIKKYVDRFIKYEKDILLADRNRKRRHRNRQPQIHKLNNKRLKKPTD  
KKQKTNNKKTWRLPKFIKKMSPASRLKFFSACIISGIKITSIIVLSIMAL  
(SEQ ID NO: 91)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
NSR1_YEAST (P27476) NUCLEAR LOCALIZATION SEQUENCE BINDING P	56	2e-07
FKB4_YEAST (Q06205) FK506-BINDING PROTEIN 4 (PEPTIDYL-PROLY	55	4e-07
GAR2_SCHPO (P41891) GAR2 PROTEIN	52	5e-06
O95367 (O95367) CBF1 INTERACTING COREPRESSOR CIR	52	5e-06
SR40_YEAST (P32583) SUPPRESSOR PROTEIN SRP40	51	6e-06
O08904 (O08904) BRAIN X-LINKED PROTEIN (BRX PROTEIN) (FRAGM	50	1e-05

CT504  
Nucleotide  
Genomic coordinates:  
Start: 13936  
Stop: 9328 (SEQ ID NO: 92)

## Amino Acid

MAHKLLFLEEEDAKEIGTSLSHPEPSFALYESETFRSVGFCKNVTDAYPKFLPRPMDINSV  
QALAVRLALIQFYKGRGWKKNMSIIDLVKDKVERNFKVDKKTSGGFIIGDGTGVGKTREL  
AAFVMSVILQEAKLLDVQKHVGPSIFGQDSDKVITAINSGVWKRHPFFIWLTCSPKLFNS  
CQQGMREVVTNSRGLRDPKFSWRKLQVPCANKPTSFKSDGKSGSMTVDVENS SVSSAKDSV  
DIRFFTLRDVKEFHSKRSSRSIGDFLTETPTILFMTYSDLRTNLEFVLKFITGGTDLDSN  
KMPIDNFEVTALLCDEFHKTQNI SDSFRKELAKTWEEDTRVLNRNIQKRANPSVSOLINR  
FKSAMSDDRNFKVKRMKSSNNKGRVTMSNYLKLLSQADAFRIFLEILKYDTFTVMASATP  
FQSNADLHMHILRKSAPAYTSIQAFKEVSSATPDAMAEHSEYVTVFLEQVIKLLNRNG  
QLVSRISISMAGVDCSTTNCKASPLQKYAIDELASYCLNARQVLIDSEKVGGHVRRRAFTKI  
IREHQEGGILEEEDVEKLVAEINSPSRKRKRKRAANDDDLYEVMENIDRRFKVVVVVRDRDVA  
HDGKTTLRISIVQDAIKTYSQKKDALSNGGGGIITSPEVDISSIDMVAQDLYDAIKKKEKP  
SKGKTDFNEDYDDGANEEDGWGEVFDDECFEKLRRQYFINTASTSVAACKGALLNIKATS  
VTDVAVKRLRTTNESKKMVMMSLEQTGDSFLKNTLTRILQTIKDESDAKYGIVDVGTFDSS  
PVANTIFSGYRLLCRAVMMASAFITISLKNKTNRRTPAHVMLVPSVPDTEPLMALAGNPI  
DSITQSIGEDSNAEITNRKLCSTRITNRGLFLVKNNKTANTNKCISAFNNTEKVDVIMLG  
PKGNTGLSLHDSNNSMYAKRYHCVLDVPYNAIAFLQTIGRTHRNQSLVSPQFLIFSTDS  
PAERRFFDSLDRKIKDSKAGTYADRYSNNSIDIAAAVMREQFIDQGLVLKTMGNIVQIVT  
ASMTKVHLMEHFSKMMRTNRGGVAFVEGLTLENGIFTEVIVLAMHIALVVIGAQNKITS  
SDDLGHALSFTSVLPHNQILSIVKSASQFVFSNLCLHLVHFKSDCDNLLPREKRVDAAS  
ALIDTLNTKNNEVTSKTNKIESDAPSLTALMLPSGPRNRKMDVFSNIMAYNNNNNGMDFDE  
DVPDNDEDEGCPLQEEENATTLALS NFPHDYDRAIKDAHQLVTVRIVGQGEKEGVIPISE  
CLDVPELDMTNLIPVVTATNVIQSLAKENPGLLFTIHNAALAHSHREGYGGSHLLGLAKK  
LSRGFINFRQFQNLFSPPKESKIMYDIFLSVKAIMARDDRYDGLCDMRMNSMMDASFLK  
LRKKPECVFITKLLDKNFRRI INDEEEETRERFGEEDDDDEEFEDEEEEAEREWG  
EEEGESAYDISVINDKNNTIGHDVIILCNRRKLTTLTKENS VFVNEHIDSFMVGNLIGAE  
GSLIQICFDNCTGEFEGLPKFCLYDSSSKDKDTIP  
(SEQ ID NO: 93)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
001940 (001940) STRAWBERRY NOTCH (SNO)	72	3e-11
Q9Y2G9 (Q9Y2G9) KIAA0963 PROTEIN	62	3e-08
O75257 (O75257) R31180_1	62	3e-08
001737 (001737) F20H11.2 PROTEIN	62	4e-08
O85862 (O85862) PROBABLY METHYLASE/HELICASE	59	2e-07
O64516 (O64516) YUP8H12R.3 PROTEIN	53	1e-05

## TaqMan Primer/Probe Sets:

5'start=2131  
5'stop=2154  
3'start=2210  
3'stop=2233  
5'primer=GGTGCACTATTGAATATCAAGGCA (residues 2123 to 2154 of SEQ ID NO: 92)  
Tm5=59.03  
3'primer=CAGTTTGCTCTAGTGACATGACCA (residues 2210 to 2233 of SEQ ID NO: 92)  
Tm3=58.06  
probe1=TTCTGTCCACCGATGCAGT (residues 2157 to 2174 of SEQ ID NO: 92)  
probe1start=2157  
probe1stop=2174  
direction1=Forward  
Tm1=69.00  
score1=1.99  
length=103



CT505  
Nucleotide  
Genomic coordinates:  
Start: 16983  
Stop: 14064 (SEQ ID NO: 94)

Amino Acid  
MDERRRDPLLYPTNRSMRFTAQITLFTVTVFVLGCFALVCAAMAYNVAKPMSVNFQAIHE  
LGMKSKLKAVQGANPEKTLEEYLEARGRHGVEDASNYPPHPALDMMNLTVKGNKWNVP  
SETKERNSRFESHDLAANRSSLVPEHHIDRLSEATIEKSNKYLDAVSGKKFRQRMVNL  
KDNIEKDDTELYDSLFGVVDIHHHSASGVSGDAPPPPPSTSEGHDEVDILAYNTGGYCS  
NPVPLKEGQCTSVCYTSRAVRVMTPFVAGGTFITHKSGEDPKPYCWGNSVPGDHIETSP  
TTGERVVKECSVHTSIVVLTDDGGWQCRPKYPTYFGGSGGTSMTACAFNPSTHKGPPPPS  
SSTPIYYDVLKKQOIRNHTEFRNSSYISKLRQSSSLAEFKIKCNDPEFLYKNPITCFCNN  
KKDVLNNDLLSQDVTKDMKFRGMYECMENPCVMPNIDPSFVTFDVSTMKCVPGVNNPQD  
SNRHAIIGDDRTPLVGTVPAMGIFLADQSKRGDQIHQORPKSSIDETAKKIALAQAPII  
TPLNLDATNTSKNVLFVPIPTVLPPLNIPHVIIIRPSSLLHRSCLAPVLNKPSSGQHRP  
ECTAPFYIEPAANVLAGNIPQKPYEHSMLATECLNRSRMVSGSVHGGSELLFSTLLSQNK  
PSSYIRTPPGGTPAPEYNSTGDQRLEEIRDFFERNFNDERRLSQTEYVIKKHARGMRTSE  
IYLKSSSWDSL MKRKEFLRHIIKKSEDTFVLKEGLLMRSYGPYAATVLARDMFDDLKLG  
KPASKTSSTLKVSNPLQYAFPTSYSVLPEEGATDDIFSVDHNRIFDSETIPSYFDCSNVT  
PGSEKLFGTSSSSSEYRVDIDDDAWGLQSFRLDHNPKSGPVVQSDPRLAFDASNISSTPE  
GATITPLSLFKKSLVEWGHKKADVQETSWFRDGVDTSEAYRRLVETSMAVRNSWFSLAW  
ENKNYYFAKNSS  
(SEQ ID NO: 95)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
CUP5_GALME (Q24998) PUPAL CUTICLE PROTEIN PCP52 PRECURSOR (	46	9e-04
CFHD_HUMAN (Q02985) COMPLEMENT FACTOR H-LIKE PROTEIN DOWN16	38	0.33
P89911 (P89911) PROTEIN KINASE	37	0.57
CAB39619 (CAB39619) AIG1-LIKE PROTEIN	37	0.57
CAB53064 (CAB53064) DJ15D12.2 (FHR-3 (FACTOR H-RELATED PROT	37	0.75
Q24552 (Q24552) TF125 PROTEIN	36	0.98

#### TaqMan Primer/Probe Sets:

5'start=1446  
5'stop=1466  
3'start=1511  
3'stop=1532  
5'primer=CAGGCATGCAATAATAGGCGA (residues 1446 to 1466 of SEQ ID NO: 94)  
Tm5=59.66  
3'primer=CTTTTAGATTGGTCGGCCAAGA (residues 1511 to 1532 of SEQ ID NO: 94)  
Tm3=59.21  
probel=GACAGGACACCGTTAGTG (residues 1468 to 1485 of SEQ ID NO: 94)  
probelstart=1468  
probelstop=1485  
direction1=Reverse  
Tm1=69.00  
score1=1.99  
length=87

## CT506

## Nucleotide

## Genomic coordinates:

Start: 25878

Stop: 25197 (SEQ ID NO: 96)

## Amino Acid

MASVFEDPADLFANMDLTGKVPTRPNILFFEGLLPNSGKEIMENRLIHKGKCGAFEADTQ  
 LAYFFPSNNEENTKKLNIGFQIKSNCLSFIRDFLNDWLEEIKDCGPYCTFSQYMDGDKE  
 IFGNSVFGQDFTIVAMDWIDKGVTFYIFVDGSDSMENMASLWMCDKLKRMNANVVKVFVD  
 NASKPKFSVCKTCRWEFPGPVSYVIEGHGMGHSDDLSCDEISEFLVQ  
 (SEQ ID NO: 97)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
AMY_BACCI (P08137) ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1) (1,	35	0.44
Q85449 (Q85449) PROTEIN 8	32	3.8
Q85439 (Q85439) COAT PROTEIN	32	3.8
VP8_RDV (P17379) OUTER CAPSID PROTEIN P8 (STRUCTURAL PROTEI	32	3.8
Q85451 (Q85451) OUTER CAPSID PROTEIN	32	3.8
Q24284 (Q24284) PLC-GAMMA D	31	6.6

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 652 of CT506: this corresponds to nucleotides 25216 to 25867 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=287

5'stop=307

3'start=351

3'stop=373

5'primer=ACGACTGGTTGGAGGAGATCA (residues 287 to 307 of SEQ ID NO: 96)

Tm5=58.18

3'primer=TGTTTCCGAAGATTTCTTTGTCC (residues 351 to 373 of SEQ ID NO: 96)

Tm3=58.28

probel=AGGACTGTGGACCATACT (residues 308 to 325 of SEQ ID NO: 96)

probelstart=308

probelstop=325

direction1=Forward

Tm1=68.99

score1=1.99

length=87

WO 01/38351

96/201

PCT/US00/28888

CT507  
Nucleotide  
Genomic coordinates:  
Start: 29077  
Stop: 28330 (SEQ ID NO: 98)

Amino Acid  
MIAIANHKKHDVSDALVGAHGAKINMLYGKSSTLSVTEAALLMFNDTALTQFAQRGYEPSI  
PTILKAALDFSLQEEEEPLVAATGLDVNKAPRSWPILNCRGLGYIASSNYPWAEHIISGDKE  
EIKRALEEHEKNANVRFDSDNCPVCLEDFSSSTNIIRTRCGHCIDKCDRLVLSTORGE  
ITRCPVCRERTSLRPDADQVKEMLVEPIVSKRMAVPDEQVSKRRRIGYNRYQFLINDV  
WTDESETV  
(SEQ ID NO: 99)

Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q9ZT52 (Q9ZT52) RING-H2 FINGER PROTEIN RHA1A	45	7e-04
CAB51420 (CAB51420) RING-H2 FINGER PROTEIN RHA1B	44	0.001
Q9ZT51 (Q9ZT51) RING-H2 FINGER PROTEIN RHA1B	44	0.001
CAB51421 (CAB51421) RING-H2 FINGER PROTEIN RHA1A-LIKE PROTEIN	44	0.001
O17099 (O17099) F42G2.5 PROTEIN	42	0.005
O16682 (O16682) ZK1240.2 PROTEIN	41	0.008

Comments:

TaqMan Primer/Probe Sets:  
5'start=257  
5'stop=279  
3'start=308  
3'stop=330  
5'primer=TCAATAAAGCACCTCGTTCTTGG (residues 257 to 279 of SEQ ID NO:98)  
Tm5=59.26  
3'primer=CCAAGGATAATTGAGGATGCAA (residues 308 to 330 of SEQ ID NO:98)  
Tm3=58.96  
probel=CAATACTGAATTGTCGCT (residues 281 to 298 of SEQ ID NO:98)  
probelstart=281  
probelstop=298  
direction1=Reverse  
Tm1=68.81  
score1=1.81  
length=74

CT508  
Nucleotide  
Genomic coordinates:  
Start: 30861  
Stop: 29073 (SEQ ID NO: 100)

Amino Acid  
MAEAAAPRYRQVLEEVLNIEPYMSFLDVFTERELALLNDIITSRNPPVPSSSFKKLDNK  
EEFRDIIYFFINNTKSDSSPICEGMTFINALTTVCKTFRGLYENIHDDFLVKYSLVLS  
MDNGFLRRETHGIKFGTGDDSRGTGFKFTSKEQAEEREKVMRRIKKLDGVLASLKKSTSS  
ARSGIVFYVEKCSSVIRFLFSRIVNITS DYVAEMKKSALEPFDISFGYKYFVDESPCV  
TKAKRLISNGNFIIGRPFSCLTSPSSVSTDFREEMNMDARSIARLNWTNEERASAYRSV  
I IKSFLSSIEEMVEEYCETTTKTVAEMAVEFVDVFIKAETIQHFQTLYSIFDTMPKFS  
AEMMDNILKNVAINAEVGSGLCGAILLWMINSRPFEEIDYNYFKICLREIMVRKKTDKLC  
DNLIVKRIVSHKNVVITDPHEVKGYVRLCVKVS CFMEDLEAFLTKNPWLKHTYFDEKGNT  
LLCYCIINKYSHTSKLVKQEKLNILKPSAKGMSPLMVCAAISSPFTTRVGIEILTNSLA  
FSFINENNENVFHAAAVATSCNFLDALAKKYKNIYDFDRSIVNARRRAMVQRP  
(SEQ ID NO: 101)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
CUL1_CAEEL (Q17389) CUL-1 PROTEIN (LIN-19 PROTEIN)	36	0.99
O23296 (O23296) HYPOTHETICAL 74.2 KD PROTEIN	35	1.7
Q57577 (Q57577) CARBON MONOXIDE DEHYDROGENASE CORRINOID/IRO	35	1.7
GIDA COXBU (P94613) GLUCOSE INHIBITED DIVISION PROTEIN A	34	3.8
Q47427 (Q47427) PLASMID P15B GENES R, S, SC, SVM1, SVMR, SV	34	3.8
Q39068 (Q39068) CYCLIN 2A PROTEIN	33	5.0

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=742  
5'stop=764  
3'start=822  
3'stop=845  
5'primer=TCCAATGGCAATTTTATTATCGG (residues 742 to 764 of SEQ ID NO:100)  
Tm5=58.75  
3'primer=GATCTAGCGTCCATGTTTATTCT (residues 822 to 845 of SEQ ID NO:100)  
Tm3=57.51  
probel=CATCCGTGTCAACTGACT (residues 797 to 814 of SEQ ID NO:100)  
probelstart=797  
probelstop=814  
direction1=Reverse  
Tm1=68.91  
score1=1.91  
length=104

CT509

Nucleotide

Genomic coordinates:

Start: 37245

Stop: 36048 (SEQ ID NO: 102)

Amino Acid

MAGTDIISSSSSGSSSSSKGGCIVSKKGKTIKGNIVFKTSIKTSSSSSEMMKKHKKRMEI  
KDMVKKCASCKKVDYSSSTLENDALRASIESTCSALNRFPEIKYGEIGEIGDVLSAIRLMA  
GCLLAKNEKSFYKFFLRGFQFDKNGFMMLSEGMRKIEKMHTKIAKKVFGGCKAAPLKEDR  
EGKIPCEFHKPSYKGEYTTPLPPTPAPVKVLPPLLPYKNVKNKPVFVFDLAVGEAKKP  
CWVHKLFSDDPPEERKRLFERHQAGRRDALMEDYGVIPNNDNEAEDTERFVSNLEYQAQM  
LELLDTANMPFPASTPVRRGRTRIVRDYDASPVSPSYSSPLHTPFDAPNVNLPNGSGRMV  
DRVRDGRNRNTRSRRTSAVMARRINQLQHQLYSSDSDF  
(SEQ ID NO: 103)

Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
001693 (001693) COSMID T08B2	39	0.073
CAB52863 (CAB52863) PUTATIVE MEMBRANE PROTEIN	36	0.37
001348 (001348) ZINC-FINGER PROTEIN KLU (KLUMPFUSS PROTEIN)	36	0.48
Q9ZDN6 (Q9ZDN6) VIRB10 PROTEIN (VIRB10)	35	0.83
Q9Y2W9 (Q9Y2W9) ENDOCRINE REGULATOR	35	1.1
GAR2_SCHPO (P41891) GAR2 PROTEIN	35	1.1

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 973 of CT509: this  
corresponds to nucleotides 36148 to 37119 of the genomic reference sequence.

CT1070  
Nucleotide  
Genomic coordinates:  
Start: 249425  
Stop: 253221 (SEQ ID NO: 104)

## Amino Acid

MASSGGFFTGIDDLFKTVIQEKEKQKNTQAPETEPKPGPSQAPDPVPDPVKPTPTNFC  
PPPPNPLPPPPPPPPKPSREERLKTSKIRLNKALSDIVEATNERVDALKENQALNTEYDK  
KDNFYQVLKCSITPSVPTAIGAHVKQVAKSSEIELAVNELDIKNKCSLVYNENESLKFF  
RDHENLILQIAVQLFSRHDNTKCVGAEICVKGNEKNKFVNKLVVKKLPNAPSSSSTVLEI  
RGATRNLLENFNKGENNTVNENKDI PPSEANLDTTKAEISHVFSTLHRLDTKRKLFFK  
GNTFYQRKPTFDNKFRTVEVIGWTESEASKQTTKSLDKPTDDNLFVLPHSFNNLADHLRL  
KFKNVLYKNSTAHGPKRNYKTQETLINPQIDSAKEYKMF AEIDKCLDVL LAIGKNDKY  
TKSTVIQYRGKFRRYLIFCYAFYALNKAHKSRAVSPLPFNFNLFSEFMYCHGPF LHSASF  
LSTLTFFVYQHMFPPMGTAAPSVSAKRLMDIDSA LMKGGKGVGVDFGSPSKTSLHTRTLV  
SFLGFAEMAMGMTMALLSGVEVRVSPALQQRISKSLEWCDSDVIFYYFTFVLFHRFSGAK  
KVSLESALRLIMGQTHAHTNKVRAAKRCRIEAAEMEGVEEEEAGLTL SYAHLGLPYSIQ  
KALGLPVPKINPLMTASSSQYNLGDVFGVEQLLAKREFPAEGETAGFLGMFDNLVKDSI  
DKYYGEGAFSDVVENVKQGMENPTPYDTSSALMTPIPKAFYEEKDVPPQEEENSTQQRYS  
LNRDVEEYLMASPMKMVFVSILDKTNQKRFMSVGDIALAVWCKRNV LKKDWEYAIK  
GNYEWLGAKMCNHL LADLVNFGILGDLKITNKLDNTDTFHRDSDRLPSVADQKKFIK  
TSLSDRKQLALVHSCVNVSTRTHVGRVTATSWAVDALRTYTRGDKDMFAALSSSLDMYHL  
GHTNSANFVPYFSRNYLCNEQENGLWGYTRRTSEKLAKELGRGLGGLNKVCVAKTELA  
AAAIAISSALDMGEVEAVMDSSKVRKIASTCLNVNAAKVSAAREKAREASIKRLLLATN  
APAAGSSRNSNRFL LKDLWGFFSDPDKRQKLIKGEAVSVLCPNTGFLHAAVPDFVIEYSF  
ESETSVIRLRRLRIKPEKQDEMVC PSTAPEANKRKLVRNNQDAVLTLDDEDNIVKYNKY  
DMVEDEEARERLRHQDKQSVIAARISKVCERKNPKKKRRLEDPELQSVDEQLIRELAAIA  
Y

(SEQ ID NO: 105)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
SSGP_VOLCA (P21997) SULFATED SURFACE GLYCOPROTEIN 185 (SSG	64	8e-09
YPRO_OWEFU (P21260) HYPOTHETICAL PROLINE-RICH PROTEIN (FRAG	62	2e-08
O13305 (O13305) PROTEASE 1	61	5e-08
P93797 (P93797) PHEROPHORIN-S PRECURSOR	61	5e-08
AAD37432 (AAD37432) M-LIKE PROTEIN PRECURSOR	59	2e-07
Q54071 (Q54071) M PROTEIN PRECURSOR, MSZW60	59	2e-07

## Comments:

## TaqMan Primer/Probe Sets:

5'start=1884  
5'stop=1904  
3'start=1931  
3'stop=1952  
5'primer=CCGAATAGAAGCAGCGGAAAT (residues 1884 to 1904 of SEQ ID NO:104)  
Tm5=59.32  
3'primer=TGGGCATAAGAGAGTGTCTAGGC (residues 1931 to 1952 of SEQ ID NO:104)  
Tm3=60.21  
probel=GAAGGTGTGGAAGAAGAA (residues 1906 to 1923 of SEQ ID NO:104)  
probelstart=1906  
probelstop=1923  
direction1=Reverse  
Tm1=68.98  
score1=1.98  
length=69

CT1071  
Nucleotide  
Genomic coordinates:  
Start: 253296  
Stop: 255120 (SEQ ID NO: 106)

Amino Acid  
MRDDTFNQETAVKLVRWYTEYDCCCPLVNRVERLLGSFGGGVDATSVRSRPALYEEDKKG  
DKCIPFRITSLIEGILLERALTTPDLAAAFDVSEKLVYCSCNNTQGNFDVSSMTIWIDG  
NNSKKYEVTCPSCTVEKISGGAESIHKKPMSLLAFFNNLVEKEAFAERIELKKLYLSLLT  
GSAAGGGGMYKDSSQQSSFNGSWTSLLFHTSKDKTRLEAEVLVSNKIKHTSRLQPRVC  
SDLLYALCSTTNNASAYAYKARNLCVIEGGEFLYFKYTI FEENGPFDSKTDLQSLVNNP  
VSETNSSALAASSSSLEDDDDCCDDDDDDDDDEDEKTKKKQPKKQTKKQKTTTSTLPPIS  
KTNHDNMLMNVLLKKGAVNGKRKMMDLSGKKGQHSKKLKTSAAGGGASSDVVAGENE  
NNPSSVSPTNNRDRKDYVLPQIEEVTIFSQHRMNNKLAESVVKHSVVINGNCLNLFV  
TQHRKKYILPHENILFCPPLVQHVGFNKFRILTGVSCEFDRIEIVFSDQSDSVVLSNNAA  
HSAILRLLSYIRENSLKRVRSTASVKGIDFVVKSQDTNIGIPLSNKEIRERQLCSASTLS  
MLAGLGK  
(SEQ ID NO: 107)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q20497 (Q20497) F47A4.2 PROTEIN	52	2e-05
KS1_HYDAT (P38978) KS1 PROTEIN PRECURSOR (HEAD-SPECIFIC PRO	50	6e-05
YIL2_YEAST (P40480) HYPOTHETICAL 123.6 KD PROTEIN IN POR2-C	48	2e-04
FKB4_SPOFR (Q26486) 46 KD FK506-BINDING NUCLEAR PROTEIN (PE	47	3e-04
Q06459 (Q06459) NUCLEOLIN	46	9e-04
NPM_XENLA (P07222) NUCLEOPHOSMIN (NPM) (NUCLEOLAR PHOSPHOPR	46	9e-04

#### Comments:

TaqMan Primer/Probe Sets:  
5'start=1007  
5'stop=1032  
3'start=1084  
3'stop=1105  
5'primer=AAACTAAGAAGAAACAACCCAAGAAA (residues 1007 to 1032 of SEQ ID NO:106)  
Tm5=57.02  
3'primer=TCAACATGTTGTCGTGATTGGT (residues 1084 to 1105 of SEQ ID NO:106)  
Tm3=57.44  
probe1=CACTTCCACCTATCAGCA (residues 1064 to 1081 of SEQ ID NO:106)  
probe1start=1064  
probe1stop=1081  
direction1=Reverse  
Tm1=68.99  
score1=1.99  
length=99

CT1072

Nucleotide

Genomic coordinates:

Start: 255074

Stop: 257477 (SEQ ID NO: 108)

Amino Acid

MLSFNPEYASWFGKMITDPGVILPVSKDVLFGSRGQSDVGIMTLDPHDLDIKITSKRIG  
 VEERLAQYNTLPMDFTAMEKELNNSRNMKESIFTGIFLDTGSAIFEDNMFNGGGSALRL  
 IRSPALNSAVFSSKNYIIKQLPTITKSLRRSQARDKQVDKTRKIVVDSFSILSAIAAQV  
 MHLTDGEMTYVPDGHCVNVVMSETNASSIYLIINDPTGSGWKIMPNNFNKTLEMRDGVID  
 RVETLVEFACKCVASSLIKRGMDLVDMQRTIRSMDFLPPASSTSNNTPRVAIMTSGSSTT  
 TGIGSLSILAEDGSTHHQIKLSEYRTGLSITENNREVSFTVEPSIDGVQAEHPLSPSILQ  
 WLPPLVKRPEVVAAAAA VVEEENGDNKPSDKDNEDKYSDDFWSNVPVTPPLITPKKWRA  
 CKINDRAMISSWKNNLVKLHKYDWTNKTTKVDYFDKMAAFVALMTFRKFQDILADNYVPP  
 QTPSQGSEYAVTMSNVATLFTDVYGFESNGNKPLFALEOLENETGIESIYVLNIIGNSPD  
 GNSVRVVRLEKEMSFLKAKQYFTEMAIPPINEKCKWTDKAPSSVKEYKYFCDLTAPISK  
 RPRKDNNDGVEHSALTYTPRCIYHTERCLVHLYSEPEKITEHVSFNKDLNILEIGKNIT  
 NQYQNTYKSI FEIVDVPIIVASMSSTKTMTVNYYIISTPSATTKFVQDPPKTGKQLLAVE  
 EVRNFKLKSVLVPPPYFRDNKRNTTLC SQITEQNCPSSEGGRFSCPSESLILKYSNLSK  
 KRALEEIAPETETSILSLAM  
 (SEQ ID NO: 109)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
TFC3_YEAST (P34111) TRANSCRIPTION FACTOR TAU 138 KD SUBUNIT	39	0.12
Q9XIC1 (Q9XIC1) F13F21.2 PROTEIN	38	0.20
P95629 (P95629) PUTA GENE	37	0.60
Q9XGN1 (Q9XGN1) TTG1 PROTEIN	35	2.3
O85475 (O85475) CELL DIVISION PROTEIN	34	3.0
Q00741 (Q00741) TAMA	34	4.0

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 932 of CT1072: this corresponds to nucleotides 256520 to 257451 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=1048

5'stop=1068

3'start=1099

3'stop=1120

5'primer=GCAGAGCATCCTCTATCCCCT (residues 1048 to1068 of SEQ ID NO:108)

Tm5=58.38

3'primer=CTGCTACCACTTCTGGCCTTTT (residues 1099 to 1120 of SEQ ID NO:108)

Tm3=58.57

probel=TTCAAGTGGTTACCTCCTC (residues 1076 to 1093 of SEQ ID NO:108)

probelstart=1076

probelstop=1093

direction1=Forward

Tm1=69.00

score1=1.99

length=73



CT1073  
Nucleotide  
Genomic coordinates:  
Start: 257551  
Stop: 259132 (SEQ ID NO: 110)

## Amino Acid

MGVQKNILVGGGGVSLLLGVVTLTGTVTEGAPAVPPFSSSSYSFTPESSVFWVEGNRVL  
SGTKKDTLINVLGKKIPYYANSIFRHDCSETRSIQWPETSPLGLNLIFCSCASHEHQHRT  
HETTEPDDLWDGSRKTTTIIILPKWWSDDVVTSLWRDNDQKCGCGQAFVSSFTSTQKEV  
QGEWLAHTNGKTSEGDINSAYLFISLQRTTLKPIITDVTEDNMMMGMRMSGTPMNPKDMT  
YFVNDFSDDIGSTPQCLVNSDILNKREEWIAVWGVADSKDLLTKHQLGEREYGSEGRRR  
NPGVEEEEEERVEEEEEVEVALPYIKKSGKLIGPRRRPLTTTTTTTTTTTTNPIVREVVE  
DFDYESFNEPEIFGSNSKLPFIRFLDQKNWRLGIMSRVSSSIANFKIEQESSKALFCLAV  
WVGDEHTPKFRLSVWKNWKPFTSAPIIVQNVGYSSDVFWHETLRSKIVDRSRDLIETKVT  
KKIGEDWANKKQTVVAMFISGIVCITVTVISIFSIVIYYKIKMPKF  
(SEQ ID NO: 111)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
G1184543 (G1184543) NAD+:PROTEIN (ADP-RIBOSYL)-TRANSFERASE,	44	0.003
O17412 (O17412) CHITINASE	41	0.015
AAF00095 (AAF00095) HISTONE ACETYLTRANSFERASE MORF	39	0.10
AAF00100 (AAF00100) HISTONE ACETYLTRANSFERASE MORF BETA	39	0.10
O15087 (O15087) KIAA0383 (FRAGMENT)	39	0.10
AAF00099 (AAF00099) HISTONE ACETYLTRANSFERASE MORF ALPHA	39	0.10

## Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 511 to 925 of CT1073: this corresponds to nucleotides 258607 to 259021 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=698  
5'stop=721  
3'start=823  
3'stop=846  
5'primer=CCATGAACCCTAAGGATATGACAT (residues 698 to 721 of SEQ ID NO:110)  
Tm5=57.04  
3'primer=GAGGTCTTTAGAGTCTGCAACACC (residues 823 to 846 of SEQ ID NO:110)  
Tm3=57.29  
probel=TTCGGACATCCTGAACAA (residues 780 to 797 of SEQ ID NO:110)  
probelstart=780  
probelStop=797  
direction1=Forward  
Tm1=69.00  
score1=1.99  
length=149

WO 01/38351

103/201

PCT/US00/28888

CT1074  
Nucleotide  
Genomic coordinates:  
Start: 274526  
Stop: 275153 (SEQ ID NO: 112)

Amino Acid  
MYIFVEGSPLTGKSSWMSKLIDTGSCGMSFLNFLRMNTSDYYNWP AEIGTEHLQLGFRET  
RVVDGMFEPVLKTFVDSWKKEQGKESLKEYLDYNGQVMEIYIAEWLRQRPLAFHVFTYTD  
EAVKSGFLNEEDLMDTATKWMAEIIREKRGNIQEIKVTPRVVFNGNVCSACFSNTRNL  
YNFGTNYNNVVHCDLLCPFARHRIVHFL  
(SEQ ID NO: 113)

# Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q9Y7X6 (Q9Y7X6) HYPOTHETICAL 75.7 KD PROTEIN	38	0.079
Q9YBC0 (Q9YBC0) 431AA LONG HYPOTHETICAL PYRUVATE DEHYDROGEN	35	0.52
O67329 (O67329) DIHYDROOROTATE DEHYDROGENASE ELECTRON TRANS	33	1.5
DPO1_BORBU (O51498) DNA POLYMERASE I (EC 2.7.7.7) (POL I)	32	4.5
YHC3_YEAST (P38742) HYPOTHETICAL 130.0 KD PROTEIN IN SNF6-S	32	4.5
O51342 (O51342) ATP-DEPENDENT CLP PROTEASE, SUBUNIT A (CLPA	31	5.9

## Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 1 to 563 of CT1074: this  
corresponds to nucleotides 274555 to 275117 of the genomic reference sequence.

CT1075  
Nucleotide  
Genomic coordinates:  
Start: 277704  
Stop: 278079 (SEQ ID NO: 114)

Amino Acid  
MWRSCISNIREMGDNKDYETRLIQRINDLESEIENKTELCEKINEQMKNLTKLYDKCFVE  
EETEFKFRKMEERVLYLKEQGIPLDPEERRTMLAEIDKSNKELDALLEENERI IKLIDEEL  
ESMK  
(SEQ ID NO: 115)

Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O44929 (O44929) MICROTUBULE BINDING PROTEIN D-CLIP-190	47	5e-05
Q13439 (Q13439) TRANS-GOLGI P230 (256 KD GOLGIN) (GOLGIN-24	43	7e-04
Q59037 (Q59037) HYPOTHETICAL PROTEIN MJ1643	43	0.001
Q25662 (Q25662) REPEAT ORGANELLAR PROTEIN	43	0.001
KINH_DROME (P17210) KINESIN HEAVY CHAIN	42	0.002
AAD29948 (AAD29948) MYOSIN HEAVY CHAIN	41	0.002

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 388 to 1 of CT1075: this corresponds to nucleotides 277689 to 278076 of the genomic reference sequence.

CT1076  
 Nucleotide  
 Genomic coordinates:  
 Start: 278221  
 Stop: 278473 (SEQ ID NO: 116)

Amino Acid  
 MQKKYDKLFEDDKRFREIEERILOOKEKGNPLDPEERLVLSADIDRSMKEIDDCLEEINH  
 IELSIDTLLEDCENLHYGLQTK  
 (SEQ ID NO: 117)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O18164 (O18164) Y6B3B.1 PROTEIN	35	0.076
P70388 (P70388) RAD50 HOMOLOG (S. CEREVISIAE)	35	0.10
Q06268 (Q06268) INTERMEDIATE FILAMENT PROTEIN	33	0.30
USO1_YEAST (P25386) INTRACELLULAR PROTEIN TRANSPORT PROTEIN	33	0.39
Q07380 (Q07380) HYPOTHETICAL 206.5 KD PROTEIN YDL058W	33	0.39
O96275 (O96275) RESA-H3 ANTIGEN	32	0.51

#### Comments:

Hit to public SBV sequence:  
 gi16007410|gb|AF178573.1: CT nucleotides 13 to 252 match nucleotides 2050 to 2289 of the  
 public sequence with a 100% homology, a score of 476 and an Evalue of 1e-137

#### TaqMan Primer/Probe Sets:

5'start=88  
 5'stop=111  
 3'start=147  
 3'stop=170  
 5'primer=AACCTCTAGACCCAGAAGAAAGA (residues 88 to 111 of SEQ ID NO:116)  
 Tm5=57.69  
 3'primer=TCCTCGAGACAATCATCAATCTCT (residues 147 to 170 of SEQ ID NO:116)  
 Tm3=57.97  
 probel=TATTGTCGGCTGATATTG (residues 116 to 133 of SEQ ID NO:116)  
 probelstart=116  
 probelstop=133  
 direction1=Forward  
 Tm1=68.98  
 score1=1.88  
 length=83

CT1077  
 Nucleotide  
 Genomic coordinates:  
 Start: 278723  
 Stop: 280976(SEQ ID NO: 118)

Amino Acid  
 MEKKTETAATTEKDPEPSVSKRSRNKEPKTTSTVYTSVKCYLSSIIKSESSRSNVTSTKE  
 RFEERCKSVSKMMVKGSLFLRLVVDECLRRYNHLEDEIDKWPDMTKDNFYVQLLRKGLDK  
 KKLKEGSTHPVEDVWNSPIVQETFLSQGEGNNPIKRHLMDFNITITYAAKQLKTCFETN  
 LRTHFRTRQQRAISGWLAENGFDKKYTKLVQHWIIGCTYKSDWVDSGDLERVKEGTKNFV  
 TLHRKHLCLVISDKNGTISYSPEEKYPIPSILNYYKFLQTEYPQNKKIQKMIVVPKHKLK  
 IHYCTFDQTTIQGICKDLGVWKMEEERHKQSEDILYKQGWYLLFDVKKIKKLRPNNWFHS  
 IQTDGEGSVLFSREVEEVEVTVSKKSKKNKKPRGDEDRRNYPPPTNAKYVVGVDPGRTNVV  
 SCSVFDTRQKRVRKHRMTAKQYYQESWMTDRRKANETYKKNNKEYKEALEEITRYDNGE  
 EIINDGNGDSTPTKKFEAYLKVVNEHYRLLWNEKGKKKYRKNNAMKVYSRKQKCISNFID  
 ELIPKRDKIEDYHIAFGDAKFACTGRGEQYASPARIFAKKIKERVGGDKRFTFVDEKYTS  
 KVCHRCNQPLNMLEKDCFSPNKKRKPPPTIVTTTTTTTTEEDEENGKWKATPLRENDRTR  
 RCSSEKTQFGYSSNRKVSTGDISMETPVPSSSTSSSFCTPTSITCVLGKGFVDRDFNASTN  
 IVHKFLGFWDKKLMEKKDKMPLKYHFIRVA  
 (SEQ ID NO: 119)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q9ZAY5 (Q9ZAY5) SURFACE PROTEIN C	46	7e-04
CYL1_HUMAN (P35663) CYLICIN I (MULTIPLE-BAND POLYPEPTIDE I)	41	0.022
O62231 (O62231) F35E2.9 PROTEIN	39	0.086
Q12080 (Q12080) P2610	39	0.11
MYS2_DICDI (P08799) MYOSIN II HEAVY CHAIN, NON MUSCLE	39	0.15
O69188 (O69188) C3-BINDING PROTEIN	38	0.33

#### Comments:

Hit to public SBV sequence:

g116007410[gb]AF178573.1: CT nucleotides 12 to 297 match nucleotides 2549 to 2833 of the public sequence with a 99% homology, a score of 551 and an Evaluate of 1e-159

CT1078  
Nucleotide  
Genomic coordinates:  
Start: 282175  
Stop: 282586 (SEQ ID NO: 120)

Amino Acid  
MGNSESRSSGIEIVHKNAGAPKRSHKTLVLSNRTERHAQIQKQIEELHHKTNKQFEQAQKV  
LDKNEERKKHQQQQQIIIPDPEERRAILAEIDKHMKEIDGFIEESERLGLLVDAEINNL  
EEKEVEEEHLLKQKED  
(SEQ ID NO: 121)

Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
ARP4_STRPY (P13050) IGA RECEPTOR PRECURSOR	45	2e-04
AAF05247 (AAF05247) ORF133	43	6e-04
MYSG_CHICK (P10587) MYOSIN HEAVY CHAIN, GIZZARD SMOOTH MUSC	43	8e-04
AAD46501 (AAD46501) LATENT NUCLEAR ANTIGEN	43	0.001
O40947 (O40947) ORF 73	43	0.001
Q9ZGM5 (Q9ZGM5) M-LIKE PROTEIN (FRAGMENT)	42	0.001

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 5 to 344 of CT1078: this corresponds to nucleotides 282190 to 282529 of the genomic reference sequence..

CT1079  
Nucleotide  
Genomic coordinates:  
Start: 286863  
Stop: 289635 (SEQ ID NO: 122)

Amino Acid  
MSSSSSSSFSFRISTYQTFLLKALAHPLVDKITQKCDTGRNQKCPHQFLADISHLIQGE  
RNGGNLFPLHPFKNQPHLEPRIVGSLHGRTLDNDIEESYCYFVKDLYNGVFSYVNGVKEL  
QGVLDKKISGSGSGESSSSRAPLIPITDVLLEYFGTLVVLPPRSKAYRVITEAVLALPF  
NEFSNNWPPTNIKAYVSRDFRMENLLAGLDHIEGEVGGSEWESI HASVVKRMVTIMRN  
KAEKKPPSTSRIFRVYVAEPVNDVTKIPIRVLSKLFSGRLAGILQKVYSYMLNLPYLL  
SSNSIDIKQGVKGITLSIPSARKLGFYLLQKDTTLQSSLSQDVADCIVSINAGIIGDDFS  
EKIRQCIEEKNKPENCCMCFCEIDKTPDFSYSSEHVARHNFFPVHAFSSSHDDKCCGAKIC  
SECIFPYIISLYEKMTGVAGVKVVDLFQCPGCKSGMLNLKGRCEFSNLCKRMILPYTST  
HCSSLFDATINRAEACFYSLEFLQYDFETARRIAHGAKDIPHVYNKVVKVNDLDRLCAL  
YCYKCVSPVVCDEFNESTDYEMVDVTPPLINLITEIVDSEYDDGPGNHMMWPAKFTCNFIA  
GSSGETPTISTCRDAVTFLGRAPRKKMAGWDDQSAVGQAI IALANWRKSGELPKNMFDDL  
EGVNAVLYRGDSFLLRAINYPVIGRSMSPSLELVKRKVNKIALIKAFFHEKRVRPDASK  
KLEWAELLVKSYLMEVLLQTPCVIHRHSFVGKTLITDELVHMRPDDATRNAYIQNL  
NAARQNAAAAASFSGSLPKPEFVPCCKERTIEWMYEKDNDVVRVNCPSCKKAIQKYGGCV  
NVFCECGTNMCWICEEKVSPADSNHCKVHRIVYSNCVRVKYALESMYGFECTMKNVEE  
GVKNYYVMENGFFFDVQEMVAKK  
(SEQ ID NO: 123)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
AAF04637 (AAF04637) HYPOTHETICAL 84.4 KD PROTEIN	213	3e-54
O97226 (O97226) PFC0175W PROTEIN	49	1e-04
Q94981 (Q94981) ARI PROTEIN	48	4e-04
CAB45785 (CAB45785) HYPOTHETICAL 262.6 KD PROTEIN	46	0.001
CAB36714 (CAB36714) HYPOTHETICAL 68.5 KD PROTEIN	45	0.002
Q9XII0 (Q9XII0) F7H1.11 PROTEIN	45	0.003

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has sequence identity to nucleotides 1 to 910 of CT1079: this corresponds to nucleotides 288641 to 289550 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=1148  
5'stop=1167  
3'start=1192  
3'stop=1211  
5'primer=TTGACAAGACGCCCGATTTT (residues 1148 to 1167 of SEQ ID NO:122)  
Tm5=59.14  
3'primer=TGGACGGGGAAGAAATTGTG (residues 1192 to 1211 of SEQ ID NO:122)  
Tm3=59.35  
probel=AGTGAACATGTGGCAAGG (residues 1174 to 1191 of SEQ ID NO:122)  
probelstart=1174  
probelstop=1191  
direction1=Forward  
Tm1=69.00  
score1=1.99  
length=64

CT510  
 Nucleotide  
 Genomic coordinates:  
 Start: 38917  
 Stop: 37381 (SEQ ID NO: 124)

Amino Acid  
 MAETVAVDEVPTCPICMGDYDSDTDCYNWSNGGMPCCRKSVHLECLFTWRFEEHMVNENH  
 LLCPCMCRAYIPVWFFFRKVEEYKYASFHSFLLSADYVNDEGVKDTLNKMSTILAPTFF  
 VPNAKGVNENEDVYMERAYTKLSFMLETLSRQEMHAFSEETFEDNHEAALMGKFKDIPPY  
 EYEGEWLKYVAPNTIDITQCLSNDDDDDEGDNNVSPSLLSGVTSFNFIEDDEDTVVVFVPP  
 EVDDNDDSESLPDLTVPPRSNNITFDTISGISSSLYDVNDDDDDDTMSLPDLNMPASAST  
 SSAPTSSAPTSTSLNINVNLCFNVDSDDDEEVIPSSSSVNQPSTSSGSSSSSSNSRKRKRP  
 RYGRDEDMSNISSESKRLCVDVKRYMCRLDNIDEEYNEIANRYLAELSALRERROETEN  
 KLGDCISRGNLFHTTVNDVIGKSLCSKKLKVKKRYASKWSANKQLIGSCLIKSASNNARL  
 DDEIAHVHSSLLNGFDTDPSEADQISSLPNL  
 (SEQ ID NO: 125)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O70567 (O70567) DENTIN SIALOPHOSPHOPROTEIN PRECURSOR	56	7e-07
P97399 (P97399) DENTIN SIALOPHOSPHOPROTEIN PRECURSOR	56	7e-07
SR40_YEAST (P32583) SUPPRESSOR PROTEIN SRP40	53	4e-06
O95815 (O95815) DENTIN PHOSPHORYN (FRAGMENT)	53	6e-06
P87736 (P87736) RING-FINGER PROTEIN (FRAGMENT)	51	2e-05
Q53653 (Q53653) CLUMPING FACTOR	51	2e-05

#### Comments:

EST confirmation of the predicted transcript:  
 An isolated EST has equence identity to nucleotides 1 to 444 of CT510: this  
 corresponds to nucleotides 37411 to 37854 of the genomic reference sequence.



CT511  
Nucleotide  
Genomic coordinates:  
Start: 40718  
Stop: 38972 (SEQ ID NO: 126)

Amino Acid  
MGGPTVITTTINTGGDHHHQQYVYHQGNKKRPVEEYNNNNYASGSTSEATTVPAYNNNNN  
NITIKTWDDVINLSITPPPPKRFKKSEVAPSPPTTRTFSNVCASKVIRQCKRQYNEWIER  
DSPYYFKGIEKSCSLEDNYDTCQQLRIGHRSIVKSSKYVHDTCFYKDKPKVGFYWPTSSC  
DEEMRFFDTRHILKELSSRNIPSSQIMDIMYMAVEVFQLPSSACERIRQKTSTLIKEVSD  
QCENWENFRKTARGCLSDLVEVPEDVKDFNTFICPWETFFEIKYGVYYIVNRGTVVKFMK  
DMNYEEFVFECVNGLSVYRKNIKGVVGVTGVCPOGLCLEMPFAGISIDDVIRCVKDSLKG  
GEYYESRDARLLYGVMVLQRMGRLEVKGVDTVAPITDSFIARKVVRSMFEKLVNMPFV  
LAETCNVITRVANEGIIINVDIKADNFVIDSISGQPKMIDLGLSYPLGYCYNDEYFRNTEE  
LIRQYIHTPPEFFRGHCLGAYSMTYSFSVMASILEDVVACSNMEGPAFNLMNMFHML  
LQSGTDTDFYQNRPSITEYALAMKHIFPFKGTVMNLFKVKK  
(SEQ ID NO: 127)

#### Top Blast Hits

Sequences producing significant alignments:

Score (bits)	E Value
41	0.017
41	0.017
41	0.017
41	0.022
41	0.022
41	0.022

Q15208 (Q15208) NDR PROTEIN KINASE  
Q56921 (Q56921) PROTEIN KINASE A  
O85239 (O85239) PROTEIN KINASE YOPO  
O68717 (O68717) PROTEIN KINASE A HOMOLOG  
CAB54949 (CAB54949) PUTATIVE TARGETED EFFECTOR PROTEIN KINA  
YPKA\_YERPS (Q05608) PROTEIN KINASE YPKA PRECURSOR (EC 2.7.1)

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=878  
5'stop=902  
3'start=981  
3'stop=1001  
5'primer=GGACTGTTGTCAAGTTTATGAAGGA (residues 878 to 902 of SEQ ID NO:126)  
Tm5=57.47  
3'primer=TGAGGACACACACCAAGTCACC (residues 981 to 1001 of SEQ ID NO:126)  
Tm3=57.98  
probel=TGAGTGTGTTAATGGCCT (residues 927 to 944 of SEQ ID NO:126)  
probelstart=927  
probelstop=944  
direction1=Reverse  
Tm1=68.98  
score1=1.98  
length=124

CT512  
 Nucleotide  
 Genomic coordinates:  
 Start: 61872  
 Stop: 60672 (SEQ ID NO: 128)

Amino Acid  
 MLSTCDLKHPSSSTDGNVLKNIHFSESIPANDIISFPSSDTEELNKDLLDSVRNQIKFGFD  
 PITETLKNCITTQTLLHSFLKSSLLTLQEKFNEWGSIQLEKGGQEMALCASLKIMGQISA  
 LIETAKEASMDNKKNNNACANCRDSKCSASLVTLFNKTIIDEKYVKQNSSSASALLANTF  
 TAGANKPPKEFITKDNAHGNSDTNYTAMSDNLICPGKYYSSDITYEVTKQAKERIKNNNK  
 KMRLATGVEMVMKELEAENNKEGGRVEVEVEGVEQQQPSTSGEEMOMEIMLPTPPPPDLE  
 SLVTEGVDDYPVFSPLPSLLSPMPASPLPSNGNSALEDGGPFAPSADIVVDKTSEIMGRT  
 PGSEVWHQRDRNSKMEIRNYGARGSGINTGRYRRNNTVL  
 (SEQ ID NO: 129)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
GARP_PLAFF (P13816) GLUTAMIC ACID-RICH PROTEIN PRECURSOR	38	0.12
YMEI_YEAST (Q03433) HYPOTHETICAL 32.0 KD PROTEIN IN CAT2-AM	36	0.36
Q9WTU0 (Q9WTU0) PHD-FINGER PROTEIN	36	0.36
O15029 (O15029) KIAA0312 (UPSTREAM REGULATORY ELEMENT BINDI	36	0.36
MML3_MYCLE (O06081) PUTATIVE MEMBRANE PROTEIN MMPL3	36	0.47
Q06166 (Q06166) MATURE PARASITE-INFECTED ERYTHROCYTE SURFAC	36	0.61

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=549  
 5'stop=570  
 3'start=626  
 3'stop=649  
 5'primer=TGCAAATAAACCACCCAAAGAG (residues 549 to 570 of SEQ ID NO:128)  
 Tm5=57.79  
 3'primer=TGCTTGGACAAATAAGGTTATCAC (residues 626 to 649 of SEQ ID NO:128)  
 Tm3=58.13  
 probel=GCACATGGCAATTCTGAT (residues 589 to 606 of SEQ ID NO:128)  
 probelstart=589  
 probelstop=606  
 direction1=Reverse  
 Tm1=68.99  
 score1=1.99  
 length=101

CT513  
Nucleotide  
Genomic coordinates:  
Start: 77506  
Stop: 76273 (SEQ ID NO: 130)

Amino Acid  
MEEHLSFNKPSPENGVVFFDFSDNTSMSNMVDNIRHRLPMDKKFSSKALLLASTPIPSDE  
QLSTKVNKAIFSHRETIVLSKALKIVVTGLIYVDGEYVDDVICLYPEKHTLNGILRYVVHL  
NMMLMDKAEDADEIRCGLIPLGRGFNREAFKFVDPVIPCAGYNILNGYHPDNQHISPSS  
TQFQVQRRCAVKQMYKQINGMFEVVVKQFSIKHNNRIFTINQVDFKGEEMKMFFALYSEEL  
LPFYSETGKLLSEKHVSKSFSQLPPHVTISVFYLRNMEEYNTLMKTDGSCFAPAIDT  
GDNFELFGMNNILVSKVCVGDALDLRRRIMEHISDAIGRNVELADNRLNPHITHCKIN  
EGVVGEWVSRLFAPCNFLCKPREIEIVFGGTFIFGRVSNNGNYVIKQPVVDYV  
(SEQ ID NO: 131)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O32743 (O32743) RECOMBINASE	37	0.23
SYA_RHILV (P24075) ALANYL-TRNA SYNTHETASE (EC 6.1.1.7) (ALA	34	1.5
O03942 (O03942) LACTOBACILLUS BACTERIOPHAGE PHIG1E COMPLETE	34	2.6
Q9X257 (Q9X257) CONSERVED HYPOTHETICAL PROTEIN	32	5.8
O22993 (O22993) CELL DIVISION PROTEIN ISOLOG	32	9.9
YK05_MYCTU (Q10851) HYPOTHETICAL 30.9 KD PROTEIN RV2005C	32	9.9

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 1 to 982 of CT513: this  
corresponds to nucleotides 76314 to 77295 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=502  
5'stop=522  
3'start=559  
3'stop=578  
5'primer=TACCACCCAGATAATGGCCAC (residues 502 to 522 of SEQ ID NO: 130)  
Tm5=58.25  
3'primer=TGCTTGACTGCGCATCTTCT (residues 559 to 578 of SEQ ID NO: 130)  
Tm3=58.35  
probel=ATCTACTCAACCACAGGT (residues 537 to 554 of SEQ ID NO: 130)  
probelstart=537  
probelstop=554  
direction1=Reverse  
Tm1=69.03  
score1=1.96  
length=77

CT514  
Nucleotide  
Genomic coordinates:  
Start: 102885  
Stop: 100046 (SEQ ID NO: 132)

## Amino Acid

MDSTSTTTIEAEKALLKEYVNENLTWEFVDRVIRHEKLMQRTDMLKLTSSRRLFSFISIIY  
SFLQDFFTTARDGVNSDEWCTQSALYHMLDGVASIIISCFRKRIDYYNKKMERLACTSIREG  
YFLVDVKTIERSRHVELLDPKKIWQRLYAEKIAPEKVVDAYNEVSKLLPDEAMANYNYRT  
GLVHLSDTLKNAKKPPTDLTMTDFDFYEKYIRSDIVLGKSNKLSGMFSENFEILPDINIK  
VPRRLERYFNVTNYSLEHNFRFPSNHIRGLIFAYFIGNIFGGAFSCVQLYLLGFTLSAA  
SACRENVLDTPFSKLKQYIKNDNKTKNSSSNEDNDGEEYYPCELQYARINSNDKNACRKS  
IVKAVKFVADRVEKASVTMMRTPIAEHESDCYMAWLSLQISKLLGRKVSASYALLFIVN  
WVAHKYKQSFNTDVGSEKYEILLKKLTVACGLTYNHKCGMVVPVIGFGSGMTNRKLRQY  
AVHCIENVIGSFISSGKRKKDIHEDPKKLEEMSLMQLSARLFKNNDVMKRGQDGKVTFAN  
EDNVQDFLEELKTKFEFVFNERRRKIHEEEYTKSLHTNLKMTFRFGVCGFQHPFPASSDKP  
TOVSLQLLKQRTFVQRETAAAVNWTRLLOFLFPSDERDNKRHQNSLPWNRLGSLNLRHF  
ISLASKFIKRSVHCERVVNDIISKFNADILPLGKDPDHFMTKAGLVIEDHARENIDNAM  
YSLCGGFNNQTTEQKLNSIRLRISAEALKNARNCVLATTFSKSYNEDRPFLPRTDEAKFV  
PIPLFGVEPLHPLLNSFIDNTANKCNDVSDFWLEESDDIFKEALVSHITLTDSSVYSTL  
VGEDEDYCDNNKSGKRIGNTLVCTLYDMMGRANYNGLHSDKPRKHDPPTWSSKNTGQSGR  
STTDFSPNSVIVLLDTENVADDYEDEEEDYEALKQSERDNVITLNNX  
(SEQ ID NO: 133)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q12537 (Q12537) GLUCOAMYLASE PRECURSOR (EC 3.2.1.3) (GLUCAN	39	0.11
AMYG_ASPAK (P23176) GLUCOAMYLASE I PRECURSOR (EC 3.2.1.3) (	38	0.32
AMYG_ASPSH (P22832) GLUCOAMYLASE PRECURSOR (EC 3.2.1.3) (GL	38	0.42
MSH6_YEAST (Q03834) MUTS PROTEIN HOMOLOG 6	36	1.2
YMM1_CAEEL (P34489) HYPOTHETICAL 81.8 KD PROTEIN K01B6.1 IN	36	1.6
AMYG_ASPNG (P04064) GLUCOAMYLASE G1 AND G2 PRECURSOR (EC 3.	35	2.1

## Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 3 to 841 of CT514: this  
corresponds to nucleotides 99939 to 100777 of the genomic reference sequence.

## TaqMan Proe Sets:

5'start=1126  
5'stop=1147  
3'start=1173  
3'stop=1194  
5'primer=TCCGTCACAATGATGAGGACAC (residues 1126 to 1147 of SEQ ID NO:132)  
Tm5=59.41  
3'primer=GGAGAGCCAGTCTGCCATGTAG (residues 1173 to 1194 of SEQ ID NO:132)  
Tm3=59.86  
probel=CGAACACGAATCAGATGG (residues 1155 to 1172 of SEQ ID NO:132)  
probelstart=1155  
probelstop=1172  
direction1=Forward  
Tm1=69.00  
score1=1.99  
length=69

CT515  
Nucleotide  
Genomic coordinates:  
Start: 114953  
Stop: 110132 (SEQ ID NO: 134)

Amino Acid  
MSRNSLRVKGKLENGGIIIPNPFDPDYVDTDAPFGMAGVKSIIIGKGFVESLLPGEISSHY  
NTFDCFKTPKKCRVGGNDFECISCRSLGGGTCVKSSRELKTEYGIEDDDDEYDGVCPPLAD  
TIFSASSAFDKHDDVDATDAAYRNVNPFPTTVEEAYLHYESGGVITGGGKKGTTYITKKRG  
CVDSSVVRKDPDLLNKDPRLEPILGCTDIVLCGGKGVGRPIHPTTFSIIDVDIDFDIS  
SMTSTMDCLCEPGYSQQRDPATNAPKEKKEGGIOEKEOGLGCPVMFRYGVVGDGTGKGC  
LCDESTQIRLEEVAGIDLDPAAKTQYAPFVEGAKLLLOITERYETLGGSTKDACLPKPG  
NDTRMSALGYSYAASFGRAPETITAFNGGHLITGGLLRESAMDAAGNWHSRIEDSDEQCK  
LTVSESVGGVVPYSGTGSVAHIWNGDALNDNGLVGAGGNGFTEHPNASLRVVPLPHSNI  
PGLGIDSIDHAVGIIASQKIFPETVHMRAGDPSPGVKTDRRDAHNDTTIETSFLKDSDKA  
GYDSYKDNPLQKLRKSHDSGICATAYVVPVSLHRVIEKPSAKNDKTVNKILPLVHYRPTA  
KRMAHTPIETIFKHSLTQERDQSFANSTLNSMMVTNSSNSFDDVTNLLLDYFFPNLNG  
EGKERSGLPINTRSIYNEPNNAKFKEIGGIILOPVTAQGAKSSTFARFSEKILSTNSPK  
IIDHYKAGSSAVFKVGEKEAYEMFAHPPTAWRIASNEGTFSSGRGLNNGIEGTGMREAE  
RVAKTLSKKPDI FAGAILTG DGVLMNGASSPLVRPMEIPASSLPEHTWFERRSPVNARGD  
PGSADNLTAINNTYDRVTGKDIRAILNSTTDIKTSFNSYAPARPFKPLAPPAGVSAQAQ  
ATSFLGVLGGFPLPIPSFLIQSVQESVSNGTGSMHGIVPLKFHEGDELWQQCEVKET  
EGALNFIPPMALFESLLRVRTLSETFIRPELIPNRFRADWGLSPHTAGHYLNGVYSPP  
CVREETGQSFGYPCSGALSQYTTMMVPKPLGPQSHSSLSKFSIKSYVEEQTRLLPANIGE  
KSIFEMQDPNSKNIFDKIGELGEKENCNCTNGLFCPKVNGGGRNKTDPIAATPSRGNRHS  
RFPLMTTLPKNDVHLSAALLRAQSGDARILNTIGETKTNGRKINLKAATENIWDISSNM  
LAPNKFCAAMRRSTAYTPYSTRQEKVPAAVLDERKGTFRNAELLGDVGMTDIVSNDILME  
DYERLPGVPAEAEIFHIIRDAAKTGQEGAKARRIVDFESSHGVTASTFNVGTFSPYVE  
GVKDIVSLYATPCFTDIDSPTISADSATINEGASIEPTDGSEVVVEVVNSNMEMLGSTA  
GSTKKRRLSISDYVDLEEDAFTINKQKATENLRVRTSSSSKYVEGGQKDMVGFYEAS  
KRVPRVMRRVHVLVPLVTPYHGGFESCAPTAQAQSACTRGVEITYADFMRPDLSTGKTTLE  
GVRVKGPPEPDDLSTLYFRSVGGPNLRKFAHHHHFGYEGLSRYYYTREKTVSVSEGLK  
DRFPFVCQSDRGFPFPPKRDGTIQPLALVDMGVLPAGALTRRTISME  
(SEQ ID NO: 135)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q24341 (Q24341) LARGE FORKED PROTEIN	37	0.96
Q24340 (Q24340) SMALL FORKED PROTEIN	37	0.96
O13788 (O13788) HYPOTHETICAL 59.6 KD PROTEIN C17G6.10 IN CH	37	0.96
CHOD_STRSQ (P12676) CHOLESTEROL OXIDASE PRECURSOR (EC 1.1.3	37	1.3
HAGA_PORGI (Q51845) HEMAGGLUTININ A PRECURSOR	36	1.6
P94986 (P94986) HYPOTHETICAL 88.5 KD PROTEIN	35	3.7

Comments:

CT516  
Nucleotide  
Genomic coordinates:  
Start: 118885  
Stop: 115402 (SEQ ID NO: 136)

## Amino Acid

MKIVQNNFTPDERSGVIHIRKPAKIEKAVFGNIAAAIDDSAAVRKDPKKRNLKNGLEPA  
SKKLAKNIERISSEELKRVTDVQDPKLLHSIMKRARQIGYDIGDDISQSA PDRDGSSS  
SSLLPIRMINIRTEELLEKGKDTIVRIHILDGILPDNVPLPFKAEIKVDLVDEKEYEGED  
GGGSSDSGSPSLFETFPFVFPAGWPPITNDPNAFSRNNGNKQQAVFKHVEVNSLADGITLS  
TKGSIFNTGNRLKISIVTEDKNKTVLFDSQVTISSPIPKITEVFACRNVALMRLDMPKAI  
NYDNVEYTPDTLNEKYVSDYPANFPRLSRQAEIASNLAAKLPRENQLSDINKPSVSFVYS  
KTNTVNTPVNLKVLNETLKNMEGNESEGYKILNATEITHLRNPSNPARTFICVSVPESE  
IEAQWKMLGWIVGFKTSSDVLTTSSGYNIVFPASKVTQSDKLFVISTDVNANTNKVVVH  
NTPSRVGCFCGSSVNFVRVDAATAPDWPGPTNGPDFFSYQLRPCIILKTDNDNREPRITAVL  
SSPATYAGERTTSLLPALNVSVGPLETEVRGGDIITPVQTALLGGEQPTFKAPAEPTKL  
YAVFPVLDSHNLVKASDNPFQPIHSITSRNKTTVLTVSDVIVNDDDDVLEDKSYHIT  
VSDPVSGSILAKENVLSSRITSRPIFIDGARDRVFSVKMEVFGGDDKGIQMPFTMDGHF  
EGQFSDMSVPSNELAIWNDPSTFTAPVRDTPATDITNKGIVYCRTLPPISNRGIRDPFM  
KQTSLVPLPTSIPEWAFADYGGEIKYPRHIFISSIRTNDTTNIVNTDTQTTEFSIENWLR  
QIDKEQERHRQLLPAPSEAYTQGEKVYAKMYMGDGVSEETLDQIVHTSNNTYVVDSEGTK  
KENLLVNKEDKKLAAILGKWGIVVFGANKYPDEPADRYTNWRNTGRLRAVGSYSQLRQPV  
APLQTRLATWPSGDPVTRLADGQFLVRLDPRCGGIGSANGFYNNNGANNEFTSSLLFAIV  
GNQDKVVSYAERVRFYMKIVARNEGKKHLKNDGVLVDRNSALHRRLLWNRTTFDHDIV  
LCVKIPQNVMSKIEPGTSSGVLVDPLVFANVASSTDREEFYKKFIDTSSGPVVIDRASVT  
SSYNISVPLNFTTCGFIVG  
(SEQ ID NO: 137)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
DNA2_YEAST (P38859) DNA REPLICATION HELICASE DNA2	37	0.68
O94534 (O94534) PUTATIVE SPINDLE POLE BODY-ASSOCIATING PROT	37	0.89
BAA84527 (BAA84527) ALP14	37	0.89
YGS4_YEAST (P46947) HYPOTHETICAL 30.5 KD PROTEIN IN SAE2-KE	36	1.5
O28907 (O28907) GTP-BINDING PROTEIN	36	2.0
Q49547 (Q49547) LMP3 PROTEIN	35	2.6

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has enquence identity to nucleotides 1 to 888 of CT516: this  
corresponds to nucleotides 115494 to 116381 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=1869  
5'stop=1892  
3'start=1984  
3'stop=2001  
5'primer=CATTACAGTATCACGTCAAGGAA (residues 1869 to 1892 of SEQ ID NO:136)  
Tm5=57.80  
3'primer=ACCAGACACGGGATCGGA (residues 1984 to 2001 of SEQ ID NO:136)  
Tm3=58.39  
probel=AGCTACCACATCACAGTC (residues 1966 to 1983 of SEQ ID NO:136)  
probelstart=1966  
probelstop=1983  
direction1=Forward  
Tm1=69.00  
score1=1.99  
length=133

CT517  
 Nucleotide  
 Genomic coordinates:  
 Start: 180036  
 Stop: 179421 (SEQ ID NO: 138)

Amino Acid  
 MEFGNLTNLDVAIIAISIAIIALIVIMVIMIVFNTTRVGRSVVANYDQMMRVPIQRRAKV  
 MSIRGERSYNTPLGKVAMKNGLSDDKMDKDVSAADVISTVTAPRTDPAGTGAENSNMTLKI  
 LNNTGVDLLINDITVRPTVIAGNIKGNTMSNTYFSSKDIKSSSSKITLIDVCSKFEDGAA  
 FEATMNIGFTSKNVIDIKDEIKKK  
 (SEQ ID NO: 139)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
SCH9_YEAST (P11792) CAMP-DEPENDENT PROTEIN KINASE SCH9 (EC 039292 (O39292) COUNTERPART OF HSV-1 GENE UL10 AND VZV GENE	35	0.45
HS78_YEAST (P32589) HEAT SHOCK PROTEIN HOMOLOG SSE1	33	1.7
G435756 (G435756) LYSOSOMAL MEMBRANE GLYCOPROTEIN LAMP-2 HO	32	2.3
LMP2_HUMAN (P13473) LYSOSOME-ASSOCIATED MEMBRANE GLYCOPROTE	32	3.0
AXO1_CHICK (P28685) AXONIN-1 PRECURSOR	31	5.1
	31	6.7

#### Comments:

EST confirmation of the predicted transcript:  
 An isolated EST has sequence identity to nucleotides 711 to 1 of CT517: this corresponds to nucleotides 179367 to 180077 of the genomic reference sequence.

#### Hit to public sequence:

gi16856160|gb|AF173992.1: CT nucleotides 1 to 615 match nucleotides 50 to 664 of the public sequence with a 100% homology, a score of 1219 and an Eval of 0.0

#### TaqMan Primer/Probe Sets:

5'start=227  
 5'stop=247  
 3'start=294  
 3'stop=311  
 5'primer=TGGCCATGAAGAATGGTCTCT (residues 227 to 247 of SEQ ID NO:138)  
 Tm5=58.19  
 3'primer=GTCCTTGGGGCTGTGACG (residues 294 to 311 of SEQ ID NO:138)  
 Tm3=58.60  
 probel=TGCTGATCTTGTCTATCTC(residues 273 to 290 of SEQ ID NO:138)  
 probelstart=273  
 probelstop=290  
 direction1=Forward  
 Tm1=69.00  
 score1=1.99  
 length=85

CT518  
Nucleotide  
Genomic coordinates:  
Start: 190743  
Stop: 188172 (SEQ ID NO: 140)

Amino Acid  
MTESKDYVLALVAETKTDEKRLNYVSEGLVAAISNLQNTPEKQRKVVISSDVFGPTWFNK  
TTEFFNSGLRLAKGHLKDAVMRSVYRDIEGVREHIIDPSWRLTETAAEELCDFTFLKQA  
PLLNLLNAFENIMDGVFSAANLVLYSTRGDTNEPSWVIDSEMLANRNNSTVADLAMGRA  
KRAIALFLGYTLCDILRWKQSIASRMKERGLDPFAAMPPHLEYGRAADMIEKRIKDFIEG  
SFSDGVTVSEEDGQSYVVPITISTVLTNMVSVIQEGFYPPKVGFSFHEALLGREIMVLLSAA  
IDAERYAVLSRTRNAKPNPLTTKLDKYVNNPHLQMPSESVTEREKEWVERERERIKTTDM  
TAENLFRDHPYLPKADIGILGPKRTPPTALQALQREYKRCNKFNDIVSPETLEYFLVNNRQ  
VMFSNYSVTRVLDPDSAARFSMYVLWNALFLCSGGLTQKTNSAVKSRLILQVFLKDMHS  
LFVCQRCESGFIKSLDTFTISLKEQSKPSMGEQELETYWKAVLDALGGGGGNNKGAENV  
NGLGELMVEILSADSGLLRGGGLGGDIGFEGKMKQKREDEEVNRMHLVDKKGIVFEAAKY  
VHVSKGFAALSFYLLYAAAATSNSPITNNFDRAVYLLARWGDLEFTHNLWGNVPTDEN  
TSSLLSFASFWALRNAVRARRNVIDASNTSFVPGRPLPLLSAFSSKMLVDNMLKNNYVKV  
ENVNREKLIWKAFREMOTSEIWKTSKKAASDRNVKAKQDLIRNASIGRLIVEPVGKT  
PISSIALFRSMKRSRSEDLKMGSSNNKYRLARDTKTATPRNPLSYTGKIVFSLDDLKNFSK  
DSYTTMKVFPPLTPLDG  
(SEQ ID NO: 141)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
CAA22155 (CAA22155) HYPOTHETICAL 46.4 KD PROTEIN	37	0.49
O95517 (O95517) DJ1170K4.1 (NOVEL PROTEIN SIMILAR TO KIAA01	35	1.9
Q43688 (Q43688) GLYCIN-RICH PROTEIN (FRAGMENT)	34	4.3
YH00 YEAST (P38800) HYPOTHETICAL 149.7 KD PROTEIN IN IRE1-K	34	5.6
P97572 (P97572) CALPAIN SMALL SUBUNIT (EC 3.4.22.17) (FRAGM	34	5.6
O48591 (O48591) GTL2 GENE	33	7.3

#### Comments:

EST confirmation of the predicted transcript:  
Nucleotides 515 to 1 of CT518: this corresponds to nucleotides 188146 to 188660  
of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=1308  
5'stop=1327  
3'start=1383  
3'stop=1404  
5'primer=AGCTGCCCGATTTTCAATGT (residues 1308 to 1327 of SEQ ID NO:140)  
Tm5=58.26  
3'primer=CCTGCTTTTCACTGCAGAGCTA (residues 1383 to 1404 of SEQ ID NO:140)  
Tm3=58.54  
probe1=TGTGCTATGGAATGCATT (residues 1329 to 1346 of SEQ ID NO:140)  
probe1start=1329  
probe1stop=1346  
direction1=Forward  
Tm1=68.99  
score1=1.88  
length=97



CT1003

Nucleotide

Genomic coordinates:

Start: 23709

Stop: 24300 (SEQ ID NO: 142)

Amino Acid

MDVSSYKSTIDYHNIEDMDDLQRATYKDRMETELVLEMAKKEGRYVRSLATMDELEVPEE

PATCYTCGYTFIRRRAPPPKRKSIFREPCAYPELLPDAPSPVRLEELVDVPEGASFFTYT

PYDDGSSTSSSSQAECEDDYPPPYDPSNPQRSQVCDYCTTRQVLSSMTDHARANLIKNLK

REKKALGLGRRNNFSY

(SEQ ID NO: 143)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
VP40_EBV (P03234) CAPSID PROTEIN P40 (VIRION STRUCTURAL PRO	35	0.46
Q9Y4G2 (Q9Y4G2) KIAA0356 PROTEIN	34	1.0
O88508 (O88508) DNA CYTOSINE-5 METHYLTRANSFERASE 3A	33	1.8
Q9Y6K1 (Q9Y6K1) DNA CYTOSINE METHYLTRANSFERASE 3 ALPHA	33	1.8
Q23804 (Q23804) SPID PRECURSOR (FRAGMENT)	33	1.8
O88799 (O88799) ZONADHESIN	32	3.9

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 736 of CT1003: this corresponds to nucleotides 23688 to 24423 of the genomic reference sequence.

CT1004

Nucleotide

Genomic coordinates:

Start: 26630

Stop: 27257 (SEQ ID NO: 144)

Amino Acid

MAPNSFQKFAPVIKTEKKEEERDEHDDPLRQIDFRDRKTLICLTANCVSRKRKAGSAHDR  
 VYKVLRYGNPYKYRRPNRTHRGLALSMDQGEVGTCLPLRPMEETEENPIDKCGVAFLYSN  
 YNEGDGMTHLYNDEEYIKKCKTIEGGTRTWVKKNRQEYFRQALETLMMSHSIKQYSNFIF  
 FKEDMEEGFVHKLHTFINMVHPKKVSVL  
 (SEQ ID NO: 145)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q9XX10 (Q9XX10) Y39A1A.22 PROTEIN	34	1.1
P70392 (P70392) RAS PROTEIN-SPECIFIC GUANINE NUCLEOTIDE-REL	32	2.5
O75381 (O75381) PEROXISOMAL MEMBRANE ANCHOR PROTEIN HSPEX14	32	4.4
Q40112 (Q40112) HYPOTHETICAL 28.4 KD PROTEIN	32	4.4
ECE1_CAVPO (P97739) ENDOTHELIN-CONVERTING ENZYME 1 (EC 3.4.	31	5.7
O70651 (O70651) GAG POLYPROTEIN	31	5.7

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 654 of CT1004: this corresponds to nucleotides 26601 to 27254 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=248

5'stop=266

3'start=291

3'stop=309

5'primer=TGGCCCTCTCAATGGATCA (residues 248 to 266 of SEQ ID NO:144)

Tm5=58.53

3'primer=CTCTTCCATGGGTCGAGA (residues 291 to 309 of SEQ ID NO:144)

Tm3=58.32

probel=AAGTAGGAACATGCCTCC (residues 272 to 289 of SEQ ID NO:144)

probelstart=272

probelstop=289

direction1=Forward

Tm1=68.97

score1=1.97

length=62

CT1005  
 Nucleotide  
 Genomic coordinates:  
 Start: 31091  
 Stop: 31961 (SEQ ID NO: 146)

Amino Acid  
 MEGEHQYLNLVREILRGVKKDDRTGTGTLSTFGPQMRFSLRDDTIPVLTTKKIFWRGVV  
 EELLWFIRGNTDAKELAKKKIHIWNANGSREFLDSRGLYDRAEGDLGPVYGFQWRHFGAE  
 YDTCSSDYTGKGIDQLANILKTLRENPDDRRMIMTAWNPMDLHLMALPPCHMTAQFYVAN  
 GELSCQLYQRSQDVGLGVFNIASYSLLTHLMASVGLKPGEFILTLGDAHIYNTHIEVL  
 KKQLCRVPRPFKLRILMAPEKIEDFTIDMEYLEGYQPHSGNLQMKMAV  
 (SEQ ID NO: 147)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
TYSY_HUMAN (P04818) THYMIDYLATE SYNTHASE (EC 2.1.1.45) (TS)	410	e-114
TYSY_MOUSE (P07607) THYMIDYLATE SYNTHASE (EC 2.1.1.45) (TS)	407	e-113
TYSY_RAT (P45352) THYMIDYLATE SYNTHASE (EC 2.1.1.45) (TS)	406	e-112
Q89940 (Q89940) THYMIDYLATE SYNTHASE	390	e-108
P90463 (P90463) ORF 70	387	e-107
DRTS_TRYBB (Q27783) BIFUNCTIONAL DIHYDROFOLATE REDUCTASE-TH	381	e-105

#### Comments:

TaqMan Primer/Probe Sets:  
 5'start=424  
 5'stop=446  
 3'start=479  
 3'stop=501  
 5'primer=ACCCTGAGAGAAAATCCAGATGA (residues 424 to 446 of SEQ ID NO:146)  
 Tm5=57.82  
 3'primer=AAGAGCCATAAGGTGAAGATCCA (residues 479 to 501 of SEQ ID NO:146)  
 Tm3=57.91  
 probel=ATGACGGCATGGAATCCT (residues 460 to 477 of SEQ ID NO:146)  
 probelstart=460  
 probelstop=477  
 direction1=Reverse  
 Tm1=68.96  
 score1=1.98  
 length=78

CT1006

Nucleotide

Genomic coordinates:

Start: 32124

Stop: 32802 (SEQ ID NO: 148)

Amino Acid

MAFNFEDSTNLFANMDLTAGTTTDPTRPNIIFFESLLPNSGIEVMKRRLVRQKCGNFEA

SGGAMSYFWLEDNAEDMENLNSGSHVKTNCLALFLQEFISNWIEETDRHGQYCTFPQYMD

GGDGSRRGGYFTSLAMKWMARDVTFVVFVDRNNTVENAASIWMYQKLLAIGAKVVVKVVDN

ASNPMFSVCNACRCKYPGPVSVVIEGHGVGHSDLCDEISGFFV

(SEQ ID NO: 149)

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 624 of CT1006: this corresponds to nucleotides 32182 to 32805 of the genomic reference sequence.

TaqMan Primer/Probe Sets:

5'start=268

5'stop=289

3'start=340

3'stop=361

5'primer=TGCTTGGCATTATTCCTTCAAG (residues 268 to 289 of SEQ ID NO:148)

Tm5=58.01

3'primer=CGTCCATGTATTGGGGAAAAGT (residues 340 to 361 of SEQ ID NO:148)

Tm3=58.72

probel=CGACATGGACAGTACTGT (residues 322 to 339 of SEQ ID NO:148)

probelstart=322

probelStop=339

direction1=Forward

Tm1=69.04

score1=1.95

length=94

CT1007  
 Nucleotide  
 Genomic coordinates:  
 Start: 32947  
 Stop: 34216 (SEQ ID NO: 150)

Amino Acid  
 MDSNTSILPPSKRPGNLNLLQVLGIIITVALIASVSSFIFYRVGKRKYYPSSSSSSSELSDV  
 DNGVEGGGGTTTTPTQSPDGGDGYVDLSPQKKAELRTRVANVIFQEVSKDQGVAFRRAM  
 NDSTDKIMEETEARINNFSEPFREATVEREVFKDDTDKNFILSTLDLTEEQFKDIVMAEV  
 KNQLENFDYEDMTRLIFDNI PETDYLTWTHFDPKKYDTYSEKVLGFSINSIERISSTFY  
 KGKKYEVTTGNVAVLVDFESETIKEKAGNSLIRNVEFIVVDEQTYKSFFPAFNQVFFSFK  
 VNKEKREVTVSINNGCVGIVANITPLTTPVGAASGHYIYGTSTAKEKTYLFVIDKYDTTE  
 FVCGLSNKSTPLMALNILFMSDTVFPSFDEAERPLTDAKAVEILGKRLGVGRYTNANIRN  
 TQ  
 (SEQ ID NO: 151)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q56711 (Q56711) HOOK-ASSOCIATED PROTEIN TYPE 3	39	0.045
O77363 (O77363) MAL3P4.5B PROTEIN	39	0.059
YLM5_CAEL (P34379) HYPOTHETICAL 49.8 KD PROTEIN D2007.5 IN	37	0.23
GBA2_CAEL (P22454) GUANINE NUCLEOTIDE-BINDING PROTEIN ALPH	36	0.51
Q9YW03 (Q9YW03) ORF MSV089 PUTATIVE NTPASE, RABBIT FIBROMA	36	0.51
P87199 (P87199) KINESIN MOTOR PROTEIN	36	0.51

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 907 of CT1007: this corresponds to nucleotides 33238 to 34144 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=648  
 5'stop=674  
 3'start=773  
 3'stop=799  
 5'primer=TGACACGTACTCTGAAAAGGTATTAGG (residues 648 to 674 of SEQ ID NO:150)  
 Tm5=58.05  
 3'primer=CCTTCTCTTTTATTGTTTCAGATTCAA (residues 773 to 799 of SEQ ID NO:150)  
 Tm3=57.41  
 probel=ATGTAGCTGTCCTCGTTG (residues 752 to 769 of SEQ ID NO:150)  
 probelstart=752  
 probelstop=769  
 direction1=Forward  
 Tm1=68.89  
 score1=1.89  
 length=152

CT1008  
 Nucleotide  
 Genomic coordinates:  
 Start: 34217  
 Stop: 35048 (SEQ ID NO: 152)

Amino Acid  
 MEGVILDKIETIAKRASPSYGSIDVGTAILRRQFMKIRGKINEETTMEKIMGTKEERED  
 TIRSIVANVIKENTVKENVTEKIRAMTDKELNDNREFMHDFGKISTGDGGTFHLFEDTPG  
 FESALKAHEYKNVPGATTPKYVSMNSLRIDAINGKIEEVYNPSPIMGIREYGTIRRGYEE  
 NAGSKELVFMTKIEKRPNNVAENLIIRVANQQYNVMRMVFIDYETKKGVSKEEMFIPYN  
 VQTKALKGRSTYFSFVRKIPDEPEGSIIHALGFY  
 (SEQ ID NO: 153)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
BAG_STRAG (P27951) IGA FC RECEPTOR PRECURSOR (BETA ANTIGEN)	41	0.007
Q99051 (Q99051) IMMUNOGLOBULIN ALPHA FC RECEPTOR PRECURSOR	41	0.007
YIBA_ECOLI (P24172) HYPOTHETICAL 31.8 KD PROTEIN IN RHSA-MT	38	0.079
Q46749 (Q46749) ORF-A1	38	0.079
Q25920 (Q25920) MATURE-PARASITE-INFECTED ERYTHROCYTE SURFAC	35	0.53
Q06166 (Q06166) MATURE PARASITE-INFECTED ERYTHROCYTE SURFAC	35	0.53

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=361  
 5'stop=384  
 3'start=429  
 3'stop=451  
 5'primer=TTTGAAAGTGCTTTAAAGGCAGAA (residues 361 to 384 of SEQ ID NO:152)  
 Tm5=58.35  
 3'primer=TCGCATCGATACGTAAACTGTTC (residues 429 to 451 of SEQ ID NO:152)  
 Tm3=57.96  
 probel=CCAGGAGCAACTACTCCA (residues 397 to 414 of SEQ ID NO:152)  
 probelstart=397  
 probelstop=414  
 direction1=Reverse  
 Tm1=69.05  
 score1=1.94  
 length=91

CT1080  
Nucleotide  
Genomic coordinates:  
Start: 291719  
Stop: 292205 (SEQ ID NO: 154)

Amino Acid  
MTSPAPSPSPKSSCTTIVNRCGFLDNNKEVVIYDTNSKFKCEPKNLELIGVLSGVSD  
NVVTQISPDQIFVGTVMKYNNWSKSGHERFSDMSNNCLDNITRPSEVIESVIKKTSSDFK  
MKYTRSLMDHTEKYFSGDQKLSKISSWCTTPIRQWVCNSV  
(SEQ ID NO: 155)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
CAB45519 (CAB45519) GIBBERELLIN 20-OXIDASE-ARABIDOPSIS THAL	32	3.0
Q52743 (Q52743) CELB	32	3.0
Q39110 (Q39110) GIBBERELLIN 20-OXIDASE	32	3.0
O60963 (O60963) I549.2	31	3.9
FLPA_ARCFU (O28192) FIBRILLARIN-LIKE PRE-RRNA PROCESSING PR	31	5.1
DPOE_YEAST (P21951) DNA POLYMERASE EPSILON, CATALYTIC SUBUN	31	6.7

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=201  
5'stop=220  
3'start=290  
3'stop=312  
5'primer=CCCCGACCAGATATTTGTGG (residues 201 to 220 of SEQ ID NO:154)  
Tm5=58.70  
3'primer=AGGGCGTGTAATATTGTCCAGAC (residues 290 to 312 of SEQ ID NO:154)  
Tm3=58.09  
probel=CGCTTCAGTGACATGAGT (residues 265 to 282 of SEQ ID NO:154)  
probelstart=265  
probelstop=282  
direction1=Reverse  
Tm1=68.90  
score1=1.90  
length=112

CT1009  
Nucleotide  
Genomic coordinates:  
Start: 35073  
Stop: 35967 (SEQ ID NO: 156)

Amino Acid  
MALQEKDITIGNVSAALRELMYSPTHMQHHDKLNFTLDRNVESSEEKIRQIVDKIRSQT  
TSDISETVNNVTNGTAFSLFEDTLEGMVKKKNIGDNLQSGDFIDGRKKLNDMKSLATGAI  
LSRQRDFVAESITGTDWLKAIMGCGIIRYTVFVNNLARSTLDNDDKAATYYNTPIYGG  
YCKMAIKDYEIPDSYSKVEAHTVEGRKMTFNIKWRGDTINNLTIIIPSVTGYLASISED  
ADVQAPLLLNCNCFIEADMSSLYMDEKKTEASFTLNLPEIEGADANAVYEICIVVV  
(SEQ ID NO: 157)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
YB75_YEAST (P38321) HYPOTHETICAL 101.2 KD PROTEIN IN FAT2-P	34	1.3
Q08281 (Q08281) CHROMOSOME XV READING FRAME ORF YOL138C	33	2.9
Q92271 (Q92271) 12.8 KBP FRAGMENT OF THE LEFT ARM OF CHROMO	33	2.9
Q26032 (Q26032) VARIANT-SPECIFIC SURFACE PROTEIN	32	3.8
O02179 (O02179) CALYX PROTEIN	31	8.6
O01394 (O01394) POLYHEDRAL CALYX PROTEIN	31	8.6

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 780 of CT1009: this corresponds to nucleotides 35166 to 35945 of the genomic reference sequence.



CT1081  
Nucleotide  
Genomic coordinates:  
Start: 292189  
Stop: 298777 (SEQ ID NO: 158)

## Amino Acid

MQLRLNFVKEEHETVVVHNPSGMTGFNIFNSSPVYFEVHNEMDALIFMAAFLKHNSLWG  
EINANMDLYTFDYAGAFLLDERWCHHEKSFSVVRQALINSYKCRRKIMQALDNNYNNKKNK  
KRKNVGGAPAFTFMSGDGEKGKEALEASFDVIGGTRGGRFGVDSTPCPHSSAMQLKLDNE  
GNYGCIACFASMFVLENPGDESSFISTDASKIGQAQAWIDERLRNNENGGEENNVFKKT  
FHMLADITQKAHETAYSNTIPLGPNGRQWNWPTHTVEPIAHEFVTHSLVNTLKNLGDRKL  
PRNFNDILYNLNPFGKMLLVFIQNCILHTGHKNENNVPRGSASGKWWTINFGVNMWT  
FQVTKCKVEKDRKISDLACMETLRLPNPGSTTVDDRIVFKGFCRGENLGSVGEVVS DIT  
QSVKNFCLMVENRKFSVDKETGFISSESIVSDPFFSLEVTGCRSNRAQDTINNGRV SARV  
MRILKSREGARVWLAKDENAIIFENVNHDTAISTDAMERAIGQHKILYYDIETTDKDFTD  
KKS VITSIGFCLCTGGDMTHGGERGVFGLVAPGSDVEKVKETIINSYDPEEKEDIMKQCP  
QVIEIFTNEFEMLLGFGKYIDKVKPHVISGWNNVAFDDPFVTRIVKHLSDHTKDMSYCV  
ADASTAESVLPRATEGGGGGETPYRLSTPQERIQLASTGIFNKLKGKFDKKTGMLKPEMT  
ADLLAGAESQANTKFKERNKLSSSNKGSAGWFQKIIGGMCSAIRLDLMKVCEKAYKESLS  
EFNLNAVLAKESSVGDVKVKNVDEVLDLHFLHLLGLFLKLLKKAQDQAKVHVYCCCKDAYLTGIV  
STSINKEGEIIFRLCMDSALTEAVVTANLATPLCIGEGAICRNMGEERADRRGVGVRHSI  
ATDTKGMVSQPIVNHVPYQIDMTSLYPMTMCQNNLCTTTTFVTHRQIMQLRDLRVLEKM  
KNKTTDSL LLLDVIDECNQIVLSEYRPIDIAVASWKNSNSNRQTPITRIEESLGLRFIEN  
LDAEKTNNKTWCTNTSPNMNVTAAGMDYFPEIVCDINMQFAAKVNDMDMHIAPASLEYMLQ  
VLPIMLIDRPYIGAHITAGKCRTELEDILSELEKDFSVEKDEEIRTHWTFKGQKQYDFCH  
SPVTQMARHIIESTGRNIRDYEGNEKFERLVLSLSDRIYRRVGAFFDSANDPAVRLWSSRLI  
NVGMLVRTWNVKTDILKGIIPQM QATYRADRVVMQNKAKEFAKMGDMKRAGLNKVGQNM  
KLGMSMYGHLALRARSSRKEFASGSANTASSISNMSATGGIGGGGTRHSVTANQITENAR  
CVFGNIGCGLOMALPGTKQTYGDTDSVFCVHNIVGDGGMPIEYDEQTGKYYYVMDIALKN  
KMAAIIPILVNSLTGKIQFVERRDAGVGMMNIAHERLAVAGLLFAKKTYHMLHFNENSAA  
FNDMIKPKSTDNNNKFASFIRKPSHADGYVVPHPNPSLILRAAEGPAGKKLKSFL EEGIH  
DEKSMEEWFTSSPTWMAMDASVINNLYASQIVGVEKGNWIDAMTSRPIEAGTEMMEAVTQ  
ANAAFTPYKKGAFVKKGITPTTKLKLGLQSLIARFLPKIEEKKSCYLDVMKNHVENFASHI  
TNPAMTITSSRVNKFDTSKESQSRPNPLALAINNHLNPSSEISLGQKFKTVTSVSSWSLSA  
EEGEVPAGYFNAGSVRW DATNMKGSVPAFSVKNLSVVPNAITSVYKMVESDKTAIKSMIA  
KNVEVLCSTANTGFSLRGALS FNTGVIVTKDVAMACIRSLNNKQMLLFVGGGKDYGED  
DDDDDEEAEEDEENGENEENKGDVTEKKIPGRSTNKDVGEETKTSEKTEGERKGSKTA  
KGKTEEIASSLSKCGKKDARDVILDRLLKATHSSCTNNEERTRVLQOYSNCTLSSYITSV  
MKLDQRVADQ MENLISQLDQIRNLSNKKRQEKGGPFKSELDAMVA AVKVKFFPVLDASRK  
LTQDHWWKCPVSIPETREEKPLMGVPFEVALNSLIGKHKCTDTCDMACCQS LYFVLLYTL  
ALKFENERLARQIGLDDSVOLMAEMLFGGDKLLAQEVLKRVKDAQDRKLVKSLPLNHNH  
DTNTIIFLFESLRFAQKPVAGMSVSEIKDAVRGLAFSTTTGTVWNYTDERFFGPLYNMDE  
LCNERVNGNCKLSFITGIYHTAAVELAAACLSCVL  
(SEQ ID NO: 159)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
DPOD_YEAST (P15436) DNA POLYMERASE DELTA LARGE CHAIN (EC 2.	52	5e-05
DPOD_SOYBN (O48901) DNA POLYMERASE DELTA CATALYTIC CHAIN (E	51	8e-05
DPOD_SCHPO (P30316) DNA POLYMERASE DELTA LARGE CHAIN (EC 2.	50	2e-04
CAB58156 (CAB58156) DNA POLYMERASE DELTA LARGE CHAIN	50	2e-04
UBF1_RAT (P25977) NUCLEOLAR TRANSCRIPTION FACTOR 1 (UPSTREA	50	2e-04
UBF1_HUMAN (P17480) NUCLEOLAR TRANSCRIPTION FACTOR 1 (UPSTR	49	4e-04

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 4 to 1205 of CT1081: this corresponds to nucleotides 297608 to 298809 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=2738

5'stop=2760

3'start=2817

3'stop=2838

5'primer=TCGTCAATCATGTTCCCTATCAA (residues 2738 to 2760 of SEQ ID NO:158)

Tm5=58.05

3'primer=TCGATGGGTCACAAAGGTAGTG (residues 2817 to 2838 of SEQ ID NO:158)

Tm3=58.56

probel=GATGACCATGTGTCAGAA (residues 2787 to 2804 of SEQ ID NO:158)

probelstart=2787

probelstop=2804

direction1=Reverse

Tm1=69.00

score1=1.99

length=101

CT1082  
 Nucleotide  
 Genomic coordinates:  
 Start: 300935  
 Stop: 305108 (SEQ ID NO: 160)

## Amino Acid

MTEQGDQGIKVRKHLHGPRGERGETGPAGAVGPAGPQGERGAIGPAGKDGAVGPAGPQGER  
 GAIGPAGKDGAVGPQGPGERGNGRPRDGAIGPAGKDGAVGPQGERGAIGPAGKDGAVGPQGERGAI  
 GPAGKDGAVGPAGPQGERGNGRPRDGAIGPAGPPGERGAIGPAGRDGAIGPAGPPGER  
 GATGIPGRDGVDSVGPQGERGEIGRPRDGAIGPAGPQGRGATGRAGKDGAVGPAGPQ  
 GEKGEAGKDGSIQPGIQQPRGETGPPGRDGTAAERGERGFPGPPGETGPPGKDGVDGSE  
 GPQKRGRTGVPVGPGEPLAGLPGRDGAIGPAGPPGERGATGLPGRNGVDGSIQPGRR  
 GATGRAGKDGAVGPAGPPGERGATGIPGRDGVDSVGPQGERGETGPAGRDGSVGPAGPH  
 GERGNGRPRDGAIGPAGPQGEKNGRPRDGAIGPAGPPGERGATGLTGSPPGRDGAIG  
 GPQGRGATGRAGKDGAVGPAGPPGERGETGPAGRDGSVGPAGPQGETGLTGSPPGRDGAIG  
 GPIGPAGPQGEKNGRPRDGAIGPAGPQGEKNGRPRDGAIGPAGPQGETGLTGSPPGRDGAIG  
 GLTGRPRDGAIGPAGPQGETGLAGLPGRDGAIGPQGEKNGRPRDGAIGPAGPQGETGLTGSPP  
 GERGETGPIGPAGPQGETGLAGLPGRDGAIGPQGEKNGRPRDGAIGPAGPQGETGLTGSPP  
 GPQGERGLKGRPGKDGEGTGPGRQGRDGMGPRLRGEKAPGNDGLEGPGRDGAIGPAGPQ  
 GPIGPQGIQGLKGIQGRPRDGMGPAGKDGIEGPRQDGTGAKGPRGLRGFQGRGTGET  
 GAQGSRGEKGDRLTGPQGRDGPGEQGLRGERGAPGPRGPRGIRGRSGPQGSNGVQ  
 GPRGPRGTGKRTGIQGLTGIEGPRGPRGIQGEKGRMGKIGHRGEKGDGDRGEQGIAGAD  
 GEKGRGLRGIRGPIGAPGKPGTEGVRGPRGVRGVPYGAQGEGLGPQGPPTGPQGPAGPQ  
 GPMGRGTGDTGMPGPPGAVGPRGEKGGRRGKNGPKGADGKDAVNIQKYSITHARAEIM  
 WEGNEIGEAYIGRSYGTDTIPVMENRIGMTNEDKKNEYCIQVMTMHSITTRGRTSGVFFV  
 VSNKTDYILLVTLMPESVSCRTDVSTNARSERVNAVRERESKSYRFIRPSDQSIGTHSR  
 SKIAVVMYPDASMSYSVDTLADVARRETTSVLLAETIHGEKDRGFYADRGTVGRMLMP  
 PTEEEELLVLQX  
 (SEQ ID NO: 161)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q14054 (Q14054) COLLAGEN TYPE VII PRECURSOR	940	0.0
CA11_CHICK (P02457) COLLAGEN ALPHA 1(I) CHAIN PRECURSOR	935	0.0
CA17_HUMAN (Q02388) COLLAGEN ALPHA 1(VII) CHAIN PRECURSOR (	933	0.0
Q63870 (Q63870) TYPE VII COLLAGEN	928	0.0
Q60444 (Q60444) TYPE VII COLLAGEN (FRAGMENT)	925	0.0
CA11_HUMAN (P02452) COLLAGEN ALPHA 1(I) CHAIN PRECURSOR	923	0.0

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 557 of CT1082: this  
 corresponds to nucleotides 304552 to 305108 of the genomic reference sequence.

CT1083

Nucleotide

Genomic coordinates:

Start: 50094

Stop: 50295(SEQ ID NO: 162)

Amino Acid

MAVTEIPCGTRNIAEEDVELELILVTAEAEVREMAAALAAAIIGAVVVQIGRVLDEVVA  
AEVELM

(SEQ ID NO: 163)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
LIPB_MYCTU (Q10404) PROBABLE LIPOATE-PROTEIN LIGASE B (EC 6	30	1.8
Q64033 (Q64033) ANTIGEN LEC-A	29	3.1
OL56_STRAT (Q07017) OLEANDOMYCIN POLYKETIDE SYNTHASE, MODUL	29	5.2
RPSD_PSEFL (P52326) RNA POLYMERASE SIGMA FACTOR RPOD (SIGMA	28	6.9
Q54540 (Q54540) RNA POLYMERASE SIGMA FACTOR	28	6.9
Q9YAS2 (Q9YAS2) 136AA LONG HYPOTHETICAL PROTEIN	28	9.0

## Comments:

TaqMan Primer/Probe Sets:

5'start=24

5'stop=45

3'start=95

3'stop=112

5'primer=TGGTACTCGGAACATTGCAGAA (residues 24 to 45 of SEQ ID NO:162)

Tm5=59.21

3'primer=CTGCTGCCATCGCCTCTC (residues 95 to 112 of SEQ ID NO:162)

Tm3=59.19

probel=TTGTAACAGCAGAAGCAG (residues 71 to 88 of SEQ ID NO:162)

probelstart=71

probelstop=88

direction1=Reverse

Tm1=68.98

score1=1.92

length=89

CT600

Nucleotide

Genomic coordinates:

Start: 1118

Stop: 491 (SEQ ID NO: 164)

Amino Acid

MHMWGVYAAILAGLTLILVVVISIVVTNIELNKKLDKKDKDAYPVESEIINLTINGVARGN  
 HENFVNGTLOTRNYGKVYVAGQGTSDSELVKKKGDIILTSLLGDGDHTLNVNKAESKELE  
 LYARVYNNTKRDIITVDSVSLSPGLNATGREFSANKFVLYFKPTVLKKNRINTLVFGATFD  
 EDIDDTNRHYLLSMRFSFGNDLFKVGEK  
 (SEQ ID NO: 165)

Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O51578 (O51578) EXODEOXYRIBONUCLEASE V, BETA CHAIN (RECB)	39	0.033
O68195 (O68195) DIOL DEHYDRATASE-REACTIVATING FACTOR LARGE	32	3.2
BACC_BACLI (O68008) BACITRACIN SYNTHETASE 3 (BA3) [INCLUDES	31	5.5
TKT_MYCPN (P75611) TRANSKETOLASE (EC 2.2.1.1) (TK)	31	5.5
Q9ZER8 (Q9ZER8) REPAC1 PROTEIN	31	7.2
CAA10001 (CAA10001) REPAC PROTEIN (FRAGMENT)	31	7.2

Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 569 of CT600: this corresponds to nucleotides 513 to 1081 of the genomic reference sequence.

CT520  
Nucleotide  
Genomic coordinates:  
Start: 209342  
Stop: 196799 (SEQ ID NO: 166)

## Amino Acid

MFKANVLNLGGGKFLES DVRDHLIKCANQMKEEPTTLRICLSNKLPEYDNRRLPLLLLNE  
GEQILVTDNLTKNGNPLVKQMGHLAVQDRVGGDGSVNPNNLLYAGCNVVEYDTVNRGNDG  
KLIMYSQPATLKDVAKSKKKGMVKVVKVPEITGDQFLDKLNERSCQENENRRMDEEGPHVG  
TGKLLRELIIMRLYEEETSSAEKLCVTPAFREFLGCGRTATDVPVFKVAFITNASLMGL  
KVIFYPTILEEERLAASDTEENVLLKSILKVQLELLSECMPIRIVERVESMIKKTVACFK  
IDIGSDNWNLP GHCKVSDTAFYPYHHAQLVGEKKNILSISNENMVTSLGVVKADRAEEMW  
CKTLESFEKKCLYLENLGSMANTDDWRRKILFSELGPEMPYRNKSLIMDQDFCTIGMCI  
KFLAEGGGLLLTKTNATLLKEKMACKGLDDSGDGDDEEDNEEGSGGKSGGSGDENNI  
NKPPPPAPKQIPPLAANVYSIINDDDKLDQIVCFKRRKHGFFLSIDIDNSPLLAMEFLLPQ  
KAMSKKNCVERVKPETKNIIRNLTGVNTIKFDTIMPFAILQIVVRYENRNKLPRTDIL  
QORLKNNTDALSKGKFAEMWQFTHKESLKPPTIEELESIPPPPTQSEEEEEAAAAAAS  
STTPDMVSSLEEGATSTSSSDENQIASLENIKKLLSIITSTFATGADKNDTIFAWTVVTL  
AERFCALYNITSHPEEYYQI IREDFEFEGGFEEKFRHMCDAINRELSIYVPKSVLEKQSV  
CRMVGAAYENSMERIKNKTNSKLCKIKYDESTMVYELNNDTFKTFDYDES DKSFGPMYEC  
APMETFQRLFAFSVKSDKEAVLADKKSEKREKLYQKQEYLRKCDNDDVSARQILNNVASN  
ESDEESDEESDDEENYGAAGGATGDY YGGDDEDDCYGFLGEFGSSDDENVPSDNASSIN  
NVQDDVFRDVFNIKTFFNFRSSLCHRQKYVSTVIVEEMEKNLC DVLTLDNSAAESGDILKE  
INRRSLMRNRWVVPFTMPVREIVKPNVNSEDTANSNNNI PFFCSCASLNNFKSDSPLSS  
NNTMSNEKCIKLP I PSSKHLKDLTVALRFNTMACERRYFSDVTAALGFVKDKVNGNIR  
SILDNRWDAIKQCKLAGKCLSSALPLGIYENVISEDNKLINTFRPRSLARLACSSGGDG  
VSDKSVNNGFFSGI WALCANQDLESVVLGSTVVDPLKPTKVFNQSLSEKELKEKRQMC  
DAANYFKDHNVS KLNIEYECFKMMEECIMRTALNGKTSNDSEFFSNLITRYGSGTNSPASR  
LWLTILETVRECFNNSLPIDWGSVLKDWGSDMLNLKAGVSNVDES GAVFELSEFLGV SAR  
AFFGKDLDTNLDADTWECLLNDNDKDWKAQVAKAYEFALKOND IRSVENFINSSNLLTNN  
NVIKKLLKIKPTSPNDVRHQI WVEDEYYPRNKSTLRSRAEWMAATEEVLKTEM SLSCLVAM  
VAMYRIMMQGESVREIATAPLRLSVDKMVPLIRCFKITSKWCSTGKGDS PKKADASIKE  
GRFYDIEEDPLHFYRFAAYVIGQVASNDIVIEEMTRKILMSFDFNGFDTSNWLQFITYRF  
SHVLMGRRSRRLSRPLSLVKNLVS VSSSLADKNSEKSNDMYEKRVGKVMKRIARLVLVKAA  
DSVRASSNDLLDCCILDVNDVSVKSLDEFRAKTRQELQETRIDTNYNLVSN SCTTAQLAA  
VEKSSRIINTNISFHNIPAGQAKVMDANEEAFIDPSLEEINKEDNSGAKQMTGKGSNRG  
RSKKS GGGGFNNAGGFYNDDSSRGSSSVVDEDSRSRTGFSQIHMDARNEEDRESGLFSYD  
GYVLNRKIMMITQONINNDIVKVISDIENFFKICVPFSKKEYALYGVTTETALSAGMDAIE  
RWNKAVEEETNKIRKECRDLDTGSVYDMNII CPGDYMSSVGE GNGCGGGSSSSGHLL  
SNNNNEANQTNEISEDQLKHEGSDCSFWFNFKVNSSEKKQKGKSVLANTGHEGRIVG  
RPLRTFIQYKKGFAETKVLTRYFSNHSDSYWSQVMPICYIKNMA LGDEDKSKKFKGKR  
PWKNFNNSNSSSNSSVKYVSIQDLEKKDSLKNVPMGYDEDL LSLYDDSLTTSTEKLENI  
KIVNDSKDAYVILGSSNQSSFDQTF SQYFTHQKISNINTYKSLGKMWN CNNGMSPKNOI  
VLLKLLLFKNLNLWIKLYERHISVLCNWGC IHPNSSKNSHFEMTKNNAPCGVTDSNPPL  
SVYHSGFLSVEDY GQLLKDTFPLMNLHRTFS AKSKDNNSSDPSEKISAASLAKAVYARE  
VLSSCLDPEGNFCTSWITNSCSVLFTPGTNIIRRGDFFNKS CYRQQDNDYCFIGKEETKK  
CPNFVSSEIEIVSILKTAVFLSTNSDGHKRVLRVINYNKDHSGLYAGIDTGCADDEDDDD  
DQGGTDKTCLLQEDSMDAKRMLISMRSVINGKSLDESSLAIKKDNFNFLAGTDKGFYLDN  
SEFNSPVQGFVAPRGTKIFKKCCDFLLNKGTGGVFARIFFTDWACIVSSSKGKNKKAI  
ESTLQIRNGGCF SERLTPSMFDNESEQELFHORYCPDFLSYDNKQNI FSEQAYKCSFLA  
NPVCPAKNMLKRAKNIRLCITNAGTALISKIMAEVEKMGNARTFISNGTAIPFRLAENTA  
CISVDNRRYFLIDGTYLLGGRLEGINLVTDMYTRCKLKAEKHVI LNSLFSTEFISAALAS  
SMEGTTMGRGLCLIEHVS YMKNTDSVSNMKNFWSMAEDQEETDENEDDDDDENEDEDEN  
EENTENTS VVKYEPVSKTAFSSSLKPPSIFIADEDI FL SILEYELAKATSDCETASSSSS  
SSSSSSSSSKHSSSSSSSNKKRKQKDDVNSTTALHALRKCYISCV DQKTGMPRMDVVYL  
LRGLMNF GGMCTAIASGDGEKAHHMVQTLCSVALNIA TKTAVVFGVTGKNNLKTTLVDLC  
KRTWFERFTNINVTALNNAGDSSSSTQANLASFAGKKGIVIIDEVGHQGSFGSKKSSSED  
DKDESASRSGNVDFGGSGGEMNSVDINEARNAYGDG GNSKIVFSNINRLMTESKLVCDQ  
EYDFISELKHENRKNACNDTKKRKRGEIEDEGVECEEIERNDGKNDENGVRIKDPINI  
SEFFARKAHWNCSSGVVSTTFKEKNIVYNMLHRGAMPFSIKDCTDSPWLNETDAVYRHCK  
KPIEYEGKFSKSEVKTALKCILGKFGSKICDNESFESIIDENCQVNNLH SWNDCKEDIDE

WNEKFMSKNKKNKQNMKIEDKVDAIMNIIQKNNGLLKWNTSFDRDGS PVLVCNPATERFS  
 EMITSSLSAQDMLEIKKYLGDNCLSTNGGVKKSVIDGNTSAPGVLIAYHCVYTGKISDDL  
 SKTNPVLLPPPKQHFAVDDAAEKALLGPTLSNINIDSIRNIKTISRKLSSIIKDPEA  
 AKLLVDRDLDFMNMERYDASLFDVVKKPSKYSFPGFTSDGSVVLSTSTSDCENVLSCLK  
 KRIEKOKMSAKNSGFIRMCMDKNLLSDEKDDSSSNSSKNTSSLPKTDDNSSDIANFLSV  
 FGENRQSSQSFASNSGGGDSNKEACFNVDTPKRRQLVSALQKHNSDGSSSIITEIAK  
 AIPQKNDVSSSITKHMLPGQFPSSLLKNMTSPQNSVMIRGIFQQGAKSSITVSPIMMSNS  
 YIFSFFVDEAMSKRLIVFPCDTTFVFENKNEDVKKIIGLLDRGMKYIHSSIMMERICKFG  
 KHGIKQRQHEFNHKKAWNDFSGHSSDNKKKDRISDVSSVLPVLMKNLIRNKVLELRDV  
 KSVSRLEENTNTFFHLYTSMSCAKAATNYGESSSSSATITEVEEDNSCDAEEQQLRRKK  
 PANYESMCNKLPSPQMCPINPKSLNTMAMNIARSRQGAQALNSMLNSVLFVEMPFVKT  
 TRFFGRDFNIKMHSPTAKNRPAINFDCIGMSLPNPDMDVVGVDKEGELIGVGSSSLTKHL  
 CDWAGSMDVDRDLMSCHHLHMLFEMALQYTECKRRLSSSLKTLKSDKTGVYVAVMLACMV  
 YQLMVSNLKYPVFLSSSSSHKRANTEDIADENQVSSLSVPMFLAMVVNKLPHALRHSTNLA  
 LPNASQKSDHSDIVKYIVMNQWGLRLNPDYLCPCNVKHVL  
 (SEQ ID NO: 167)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
SR40_YEAST (P32583) SUPPRESSOR PROTEIN SRP40	54	2e-05
Q9Z1T1 (Q9Z1T1) AP-3 COMPLEX BETA3A SUBUNIT	54	2e-05
AAD56625 (AAD56625) NUCLEOLIN-RELATED PROTEIN NRP	52	7e-05
P90493 (P90493) HERPES SIMPLEX VIRUS TYPE 2 (STRAIN HG52),	52	9e-05
P70475 (P70475) NEURAL ZINC FINGER TRANSCRIPTION FACTOR 1 (	50	3e-04
O95815 (O95815) DENTIN PHOSPHORYN (FRAGMENT)	50	3e-04

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1064 to 1 of CT520: this corresponds to nucleotides 196778 to 197841 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=7375  
 5'stop=7397  
 3'start=7450  
 3'stop=7472  
 5'primer=GATGACGATCAAGGTGGTACAGA (residues 7375 to 7397 of SEQ ID NO:166)  
 Tm5=57.66  
 3'primer=CCATTAATGACAGACCGCATAGA (residues 7450 to 7472 of SEQ ID NO:166)  
 Tm3=57.78  
 probel=CAATGGATGCTAAGAGGA (residues 7424 to 7441 of SEQ ID NO:166)  
 probelstart=7424  
 probelstop=7441  
 direction1=Reverse  
 Tm1=68.98  
 score1=1.98  
 length=98

CT601  
Nucleotide  
Genomic coordinates:  
Start: 1511  
Stop: 1196 (SEQ ID NO: 168)

Amino Acid  
MYLSHIRQTPLVEERRALTFKMYHHNNNNQHSFVNCQCRRTSSSINCSSCSRETFNSVKA  
IQYFNKTSRNNTAHHFKMPASKDRNYSSFYEAETAVAAHNISQW  
(SEQ ID NO: 169)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
YT66_CAEEL (Q11082) PROBABLE G PROTEIN-COUPLED RECEPTOR B05	32	1.4
O00885 (O00885) MAP KINASE KINASE PROTEIN DDMEK1	30	4.2
CHD1_HUMAN (O14646) CHROMODOMAIN-HELICASE-DNA-BINDING PROTE	30	4.2
O97292 (O97292) PFC0965W PROTEIN	30	4.2
O96226 (O96226) SER/THR PROTEIN KINASE	30	5.5
O96563 (O96563) C-13 ANTIGEN (FRAGMENT)	30	5.5

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has sequence identity to nucleotides 353 to 3 of CT601: this corresponds to nucleotides 1184 to 1534 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=96  
5'stop=116  
3'start=142  
3'stop=162  
5'primer=CTTTGTGAATTGCCAGTGCAG (residues 96 to 116 of SEQ ID NO:168)  
Tm5=58.03  
3'primer=AGTTTCACGGGAACAGCTTGA (residues 142 to 162 of SEQ ID NO:168)  
Tm3=58.43  
probel=TCTTCCTCCATCAACTGT (residues 124 to 141 of SEQ ID NO:168)  
probelstart=124  
probelstop=141  
direction1=Reverse  
Tm1=68.98  
score1=1.98  
length=67



CT521  
 Nucleotide  
 Genomic coordinates:  
 Start: 272387  
 Stop: 268691 (SEQ ID NO: 170)

Amino Acid  
 MSKSSSTVKSASFNSLMENAPSSKIELLEDGWTKKAAAADTDPTAKPTGLSISLMDI  
 SGSMGSVKSAVADSCSGIMATLNVIAPGIQNAIVYYNDFDKHSIESGPVVRAPDCSEWEG  
 GDFVKHMRKTEVCGGGGGSEALHSSLMYVFNNMIPAFKKMHGITRDEKFPILIFVFTDE  
 DVRIANSDTGKLCANSYDSETAPEEEFIMKTWGQKPLTILDMRKALVENDCWLRIINFSR  
 CSGSNQSELQCEDVINFSGYDNNRWOLFESFDRRSCNVRKNIATFIMRQISLKFKNLNDQ  
 FSAFPILREINQEELNVFIESEGRSEPAGFEKYGDAQRESFKSRVLNMAPLDFGRVVQGG  
 GRYNNHKRSVFLNCAYDSAFCCSKQTFNPQQQQQQQSSSGGGGISKLAVVTQRAQSITG  
 GGNAASTLALHMNACFQSLDDFGIDHTNLCDCKGCTKLMAVEATSDQGRKTKLSRKYAR  
 VHWAKMFAEKLFKMMIKEQSMYACSAVPDEIGAIYAFVTGNNAGVCSRVTILSDLGTE  
 CGNKA EYAFLEKGGKHKMSASYDALQVINNTDLTPEQSSMFMWFYVPNDAL EEA GKFHQS  
 FSFSNSYTGGLLSLDEYKRFEFGQCFDFIKKLVSLKITRNVEDVLL ETSKTSNRYFAI  
 PVFCGSD DQKEVLREELASDLFGGREDVAEMMFIDLETVIQKLGTL YDVRLSLPEGGYAA  
 IKSVC AAASWAASCEVPSNTSNMILSI AKMAFTKYYQE QNSSSETDLDIILPSILEGTAD  
 GEIENNL SGVFLRCLITWANKIGVDKNFTNKLEHFLALRILTKAGDSKIGEKYETFPVR  
 RDLSEKDLKYICKRCGVKSLKMEYDNDEKLC LRCKGN YRMGKPMVYHWDNKLTRDPRAK  
 TASPTTLNLLNAKKIDDKVKEMASDIIGALNLPPTDKDNEI AVSAAAKAVGILYKGTCLL  
 YKLLNEGNIDI PVAVC VECDCCKSKYMMSTLGPDKPQNRKCPWCRYANKLVAMGRGGKKL  
 LMDLIECGAPSLAMVEEAIRTS GDMYEEELGEGEEFYIIDYFLK LKNTAIAEGNKLQNN  
 NKRPA PQLVTS PSSPPKKMRSDLPDSLLAAIGECAIETKEKTTVNLIGLGEVKVVENVGP  
 NDLDGKDPFISLQ EYCSWDKFNSLFVNPWLGYRLDEQWDDWNTFLIHVKKNDVWKFLCNK  
 TSPFSVVVMNDGSGLLNVDNVNVLVRQKICV  
 (SEQ ID NO: 171)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
AAF04634 (AAF04634) HYPOTHETICAL 45.2 KD PROTEIN	76	1e-12
Q9Y1T1 (Q9Y1T1) DNA POLYMERASE ALPHA CATALYTIC SUBUNIT (EC	36	1.3
P91805 (P91805) ARYLPHORIN GENE-SPECIFIC BINDING PROTEIN-2	34	6.4
O45322 (O45322) DY3.5 PROTEIN	34	6.4

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 895 of CT521: this corresponds to nucleotides 268736 to 269630 of the genomic reference sequence.

CT602  
Nucleotide  
Genomic coordinates:  
Start: 2996  
Stop: 2702 (SEQ ID NO: 172)

Amino Acid  
MDILEDIYKSAITLVLSPEFVNDVKQEASQVVEGLIPSIREAVFRRLLLEERKKHEDEV  
GDVEDKRQAVIDKANTMITTMAAEYLESVDILEEFGF  
(SEQ ID NO: 173)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
CRCM_HUMAN (P23508) COLORECTAL MUTANT CANCER PROTEIN (MCC P	36	0.055
ACVS_PENCH (P19787) DELTA-(L-ALPHA-AMINOADIPYL)-L-CYSTEINYL	33	0.48
ACVT_PENCH (P26046) DELTA-(L-ALPHA-AMINOADIPYL)-L-CYSTEINYL	33	0.48
Q9XJC5 (Q9XJC5) HYPOTHETICAL 26.4 KD PROTEIN	32	0.62
NOSZ_ACHCY (P94127) NITROUS-OXIDE REDUCTASE PRECURSOR (EC 1	32	0.82
EZRI_BOVIN (P31976) EZRIN (P81) (CYTOVILLIN) (VILLIN-2)	32	1.1

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=121  
5'stop=144  
3'start=200  
3'stop=222  
5'primer=AGAGAAGCTGTCTTTAGACGGCTT (residues 121 to 144 of SEQ ID NO:172)  
Tm5=58.14  
3'primer=TGCCTTGTCTATCACTGCTTGTC (residues 200 to 222 of SEQ ID NO:172)  
Tm3=58.14  
probel=AACACGAAGACGAGGTGG (residues 164 to 181 of SEQ ID NO:172)  
probelstart=164  
probelstop=181  
direction1=Forward  
Tm1=69.14  
score1=1.85  
length=102

CT522  
Nucleotide  
Genomic coordinates:  
Start: 276736  
Stop: 275206 (SEQ ID NO: 174)

Amino Acid  
MASGFAIKGIVKNYRRIPSIIIESIKSIRRSELAEGVYIVSLHKNTPKHEVDEIVNKIRLS  
AGNPCLEKTSFLQHHSQMRNFYTRKGAESESDWLKRLPEDLRNINNIVKREALPHDKSF  
TFSPLYRILTDRLFNAAIHNCKYIIVTADLLMCGGITNNKVEKKLLSMGSILGGESMVPL  
HDIHRLSYKGLRIENPIVGSCHDQCLVVPVSMGLKIFSSNMYPTFKNFDQCMALFLNAV  
VTHSAEKMDGKHERNKVIHMPNEVYLDAARRKYLEEKLEETNKLDAIDEEAREEYGNIG  
RIGDKSTCLVFALSARDFFLNRFNEDTPLYSGTERGIRFMCNMYCTMRDEGGFRPRIM  
SAYGPTSYPIIFNTLYDQFNQYPCVSGVLSFIGDDQLAPEPESLVDIVVRSIKNPST  
RIFSGDCETVYQDGRDVDVGEGKKNQKFNREERTILNVLRIIKAYNEERTKEDEDEEEEE  
EEEEEQQTAAATVTVESDWDSLGERGENWV  
(SEQ ID NO: 175)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O61310 (O61310) TSJ5	44	0.003
CINA_ELEEL (P02719) SODIUM CHANNEL PROTEIN (NA <sup>+</sup> CHANNEL)	42	0.009
MYC_BRARE (P52160) MYC PROTEIN (C-MYC)	41	0.019
Q9Y0C6 (Q9Y0C6) J5 PROTEIN (FRAGMENT)	41	0.019
O35788 (O35788) CYCLIC NUCLEOTIDE-GATED CHANNEL BETA SUBUNIT	41	0.025
MYC_CARAU (P49709) MYC PROTEIN (C-MYC)	40	0.033

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has sequence identity to nucleotides 374 to 1 of CT522: this corresponds to nucleotides 275340 to 275713 of the genomic reference sequence.

#### Hit to public SBV sequence:

gi|6007410|gb|AF178573.1: CT nucleotides 1 to 552 match nucleotides 555 to 6 of the public sequence with a 99% homology, a score of 1063 and an Evalue of 0.0

#### TaqMan Primer/Probe Sets:

5'start=682  
5'stop=703  
3'start=767  
3'stop=787  
5'primer=AATTTTGATCAATGCATGGCAT (residues 682 to 703 of SEQ ID NO:174)  
Tm5=57.93  
3'primer=CGTTTGGCATATGGATGACCT (residues 767 to 787 of SEQ ID NO:174)  
Tm3=58.39  
probel=CAGTTGTTACACATTCGG (residues 716 to 733 of SEQ ID NO:174)  
probelstart=716  
probelstop=733  
direction1=Forward  
Tm1=68.99  
score1=1.99  
length=106

WO 01/38351

137/201

PCT/US00/28888

CT603

Nucleotide

Genomic coordinates:

Start: 24906

Stop: 24660 (SEQ ID NO: 176)

Amino Acid

MTCPEISKHISGTDRRFWNTADPGGLSYPFNPLFTLHLHLKNFSKIFSAHSSLGGGPLTR

PYVKFEGWTAGSTQRQITERS

(SEQ ID NO: 177)

Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
CAB49803 (CAB49803) HYPOTHETICAL 69.0 KD PROTEIN	31	1.8
O74312 (O74312) PUTATIVE TRANSMEMBRANE	29	5.2

Comments:

Hit to public SBV sequence:

g116165655|gb|AF099142.1: CT nucleotides 118 to 231 match nucleotides 5652 to 5539 of the public sequence with a 90% homology, a score of 139 and an Evalue of 5e-36

CT523  
 Nucleotide  
 Genomic coordinates:  
 Start: 281865  
 Stop: 281127 (SEQ ID NO: 178)

Amino Acid  
 MVSTRSMEAKAAAAAKEVSPTTSKRKAEDLTEGTEEEESVETHPPSKLPRVDEDEVY  
 IDENVGDVQILASSIEVARMERERLAEAMVRDIKIEEKAATEARKEIASRLIYKEMVY  
 LLPQLENMTNRLRPRSLRLHNEMTITDRTFSDLQIFNKVTFEPILTIDIAFLAREKSRVE  
 GSRFYNDMKIGPITAYKLNLMCNKFIESVVQKVKAESPFVEVSVSSELEGSPFWDFKQR  
 IVKHT  
 (SEQ ID NO: 179)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q23915 (Q23915) PROTEIN KINASE	43	0.001
O77819 (O77819) CORNEAL EPITHELIAL RHO-ASSOCIATED-SER/THR K	43	0.002
Q13464 (Q13464) RHO-ASSOCIATED, COILED-COIL CONTAINING PROT	43	0.002
P70336 (P70336) RHO-ASSOCIATED COILED-COIL FORMING KINASE 2	40	0.017
P70335 (P70335) RHO-ASSOCIATED COILED-COIL FORMING KINASE 1	39	0.029
Q63644 (Q63644) RHO-ASSOCIATED KINASE BETA	39	0.029

#### Comments:

EST confirmation of the predicted transcript:  
 An isolated EST has equence identity to nucleotides 1 to 712 of CT523: this  
 corresponds to nucleotides 281143 to 281854 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=404  
 5'stop=425  
 3'start=502  
 3'stop=522  
 5'primer=GATCACTTCTCAGGCACAACGA (residues 404 to 425 of SEQ ID NO:178)  
 Tm5=58.93  
 3'primer=ACGGGCAAGGAAAGCAATATC (residues 502 to 522 of SEQ ID NO:178)  
 Tm3=59.06  
 probel=AATGACCATTACAGACCG (residues 426 to 443 of SEQ ID NO:178)  
 probelstart=426  
 probelStop=443  
 direction1=Reverse  
 Tm1=68.98  
 score1=1.98  
 length=119

CT604  
Nucleotide  
Genomic coordinates:  
Start: 50300  
Stop: 49079 (SEQ ID NO: 180)

Amino Acid  
MSHINSTSAATTSSNTLPICTTTAPMIAAARAAAIAASRTSASAVTSINSNSTSSSAMFRV  
PQGISVTAMPPVPALTSLTESTGTRMSSTPNVDVIPVPGPKNKSKKKDSKRKKQNQNGN  
RSSDEDEPSLVIDDGSGRQSKNKKYSWVTSLATTTAERNNDTLAPPRPFLPTPEEGNMSE  
IDAGLSNPVTRQITGEVYSAALTSGVGDNGLYPSHFTVADTSYGDCETPIPGPAFVLDDG  
TVSRGTSLLHREEAEFLNDGSKVIHTVKPRNSKYSNIQRAASCMAYAVDLLNNHNITSQ  
FDFMAMTAAARQRCGEMAKFFEKRDKDIGEYRNKVQYNRGIFTRTTEMNKRAKIILEQ  
QQRREAAAAAATGATAPIPTTSAAGVGATSSATTNSLEYQEIRYQ  
(SEQ ID NO: 181)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
AMYH_YEAST (P08640) GLUCOAMYLASE S1/S2 PRECURSOR (EC 3.2.1.	55	8e-07
AAC49609 (AAC49609) GLUCOAMYLASE	55	8e-07
O39781 (O39781) MEMBRANE GLYCOPROTEIN	53	4e-06
O39782 (O39782) MEMBRANE GLYCOPROTEIN	50	3e-05
Q9Y075 (Q9Y075) PROTEOPHOSPHOGLYCAN (FRAGMENT)	48	8e-05
Q14888 (Q14888) MUCIN (FRAGMENT)	48	8e-05

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has sequence identity to nucleotides 1 to 666 of CT604: this corresponds to nucleotides 49128 to 49792 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=572  
5'stop=592  
3'start=640  
3'stop=658  
5'primer=GCCAAATCACCGGAGAAGTTT (residues 572 to 592 of SEQ ID NO:180)  
Tm5=59.06  
3'primer=CAGCAACCGTGAAGTGGGA (residues 640 to 658 of SEQ ID NO:180)  
Tm3=59.24  
probel=CTCACTTCTGGAGTTGGA (residues 604 to 621 of SEQ ID NO:180)  
probelstart=604  
probelstop=621  
direction1=Forward  
Tm1=69.03  
score1=1.96  
length=87

CT524  
Nucleotide  
Genomic coordinates:  
Start: 291298  
Stop: 289681 (SEQ ID NO: 182)

Amino Acid  
MEDFKQLKVKNGICLSGENTENYERVLLTFKSVKSVRRSELKEGHFIVRLRDKEVLHIKN  
GNERLRQLTGDPTLQIGLKYTSSLPKQGSFLEDEDPNYGKKWNESLPSPFQEMNKIVEEK  
ALVNDKNFKFSPLYRIIHERLSNAAVKKCDYMIITDFLVGC GFSPRNCTRTLKNMEQVL  
VQHGGTSSRVSVYDIDRLTYNGLSIANPIVGSFSNMCLIVPMDKLGLLFYNSTHPSAKS  
IGNYMSCLFNAAVVYTLKSNQKLDNFEKEIRFAKNEVNLLVSERSVLEEKLKESKKLYA  
ASEEQRISLRDVHKKSSIASSRYDGGACLVFAFSDRDFSLLCRTNGNGSFYSATEEGIRY  
VSSDDYRKRDVDERRPRLVMSITGSDAPICIRDSIRNHFNHFIASGKGNEISFIDPPNE  
RLLMEMVREVTGSDIKIFMDNGKVYQDGV EIKVIDPSSKEGKDIIKKEETLP EEEKRLR  
RERRMIFNTVKAITYNEERGE EEEVATSSGGTKRKREEKEGDYVALLNKACKEIKVC  
(SEQ ID NO: 183)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q57105 (Q57105) FESMID CLONE 31, COMPLETE SEQUENCE	39	0.10
O94986 (O94986) KIAA0912 PROTEIN (FRAGMENT)	39	0.10
O94944 (O94944) KIAA0866 PROTEIN	38	0.23
MYST_HUMAN (P35749) MYOSIN HEAVY CHAIN, SMOOTH MUSCLE ISOFO	38	0.23
YDH6_SCHPO (Q92351) HYPOTHETICAL 140.8 KD PROTEIN C6G9.06C	38	0.23
Q22705 (Q22705) T23G7.3 PROTEIN	38	0.23

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 1642 to 1 of CT524: this  
corresponds to nucleotides 289650 to 291291 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=723  
5'stop=746  
3'start=828  
3'stop=852  
5'primer=TGGAAATTACATGTCATGCCTTTT (residues 723 to 746 of SEQ ID NO:182)  
Tm5=58.47  
3'primer=TTGCTGCTAACTAGAAGGTTGACTTCA (residues 828 to 852 of SEQ ID NO:182)  
Tm3=57.59  
probel=TGCTGCAGTTGTATACAC (residues 750 to 767 of SEQ ID NO:182)  
probelstart=750  
probelstop=767  
direction1=Forward  
Tm1=68.96  
score1=1.96  
length=130

CT605  
Nucleotide  
Genomic coordinates:  
Start: 51809  
Stop: 50423 (SEQ ID NO: 184)

Amino Acid  
MDSSASVVFMRFPAPGEETALPPRRATPGSVAYDLFPSEEMDIEPMGLAKISTGYGIDKF  
PDGCGYQIVSRSGMTWKNNTSVPTGTIDVDYRGELKVILRNHSAEKSVPIRKGTSIAQLI  
FLRYCDVEEEQIVYINETTGERTIIDSSSKKDNKNQARSVRGTGGFGSTDNPNFTETTVS  
RNQQEENKKEELEEGEIVEMEGFIDIPFLEGFENILAEQSNETGVTPNTNQDVEEKDTK  
NIDVVRELEAEFSSGIGSGSMDSSDSSSSSSSSDSSDSSDSSDSESSDDSEGGDNKVR  
IRRHQYHRRQLSYSDDVNGGGRNSEKMEMDRVTHIKTEHIKREDEPRYEERERYIHPRRM  
QVPKDYCEQYEHYDAPAAAHHRHHQHRHQHQRHFNQPRSNSSDVTAYVNENSPTRPC  
RDRNSRFSERPNGGYNRINSRYTTFDPYRYGARRGRGGVY  
(SEQ ID NO: 185)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q89662 (Q89662) COMPLETE GENOME	90	2e-17
Q86612 (Q86612) ORF2	89	8e-17
DUT_CHVP1 (O41033) DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDOH	88	1e-16
DUT_YEAST (P33317) DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDOH	87	2e-16
DUT_LYCES (P32518) DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDOH	83	3e-15
DUT_HUMAN (P33316) DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEOTIDOH	83	5e-15

Comments:



## CT606

## Nucleotide

Genomic coordinates:

Start: 65024

Stop: 64010 (SEQ ID NO: 186)

## Amino Acid

MSSSQGLNNNMCTTEILLPKCTSSSSLSLEESVDYLEKDFEELGIPLVEGKEVLLLEFAYKI  
 LNKRDITRVIGDEQGDVCSVFFLRFGKKKTFNPQTKMWLVKLANAIALSMGVVPEPACTC  
 SRMMTTAKKIPVPESYKNVNRNIQKFEDVHYIDINFQSFVREQIGLSVLGKNDVQKKKKE  
 ETPFFAFPNKSKIGGECIEDLKYDSESVSIIRDVFNLLGEMPTEDVKTSRSCINPSHNDT  
 NPSMRLVFRPMYWRNSKLVMDKLSKEQDSALIEKYMGGEHQHCCIIGGRNVLLYCITALCF  
 SSDCGFKKMLTNDDEIKQLIWYLVLLFFHIIICPIIQSK  
 (SEQ ID NO: 188)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
YNHE_ECOLI (P77522) HYPOTHETICAL 56.3 KD PROTEIN IN LPP-ARO	35	0.69
TTK_MOUSE (P35761) DUAL SPECIFICITY PROTEIN KINASE TTK (EC	34	1.2
Q56333 (Q56333) FLIL	34	1.2
CAB49508 (CAB49508) HYPOTHETICAL 32.8 KD PROTEIN	33	3.5
BAA85006 (BAA85006) ORF1P	32	6.0
BAA85071 (BAA85071) ORF1S	32	6.0

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 531 to 1 of CT606: this corresponds to nucleotides 64063 to 64593 of the genomic reference sequence.

## TagMan Proe Sets:

5'start=546

5'stop=567

3'start=655

3'stop=677

5'primer=CCCTTTCTTTGCACCCTTTAAT (residues 546 to 567 of SEQ ID NO:186)

Tm5=57.43

3'primer=ACATCCTCAGTAGGCATTTACAC (residues 655 to 677 of SEQ ID NO:186)

Tm3=58.14

probel=GTATGATTCTGAGTCTGT (residues 606 to 623 of SEQ ID NO:186)

probelstart=606

probelstop=623

direction1=Reverse

Tm1=68.79

score1=1.79

length=132

CT607  
Nucleotide  
Genomic coordinates:  
Start: 68659  
Stop: 65032 (SEQ ID NO: 188)

## Amino Acid

MGVPEAKKVYENAYGAQNGRVIKEKTGYEDCYDDEDDDEDYCSGEEDCTTSSLLKATSLAN  
INSKNFLDFGRGKKSSSSSPTCDYTLDMDVLPYNSDLVMLGRQIATTMLKGQKNMGQM  
ILFINTTNQQIIDVLHDGFNVIREEDTMHSRMQNKHHIYENFYCRDEKKVISEFFSRKYK  
HEKIKARIERVPIIIPSSQEEVDWLTEPPIEDMMMAPPVSNHKMDDYEGLDYWINKHTDV  
MKKRKFLTNLFLFRNVPTTSFNSSPTAVLKSRFKDAFFASQMEGVILYYAFRMRVMKNL  
LKSKNLKGRYTVLFTDGKAPAIKMMTRAKRQIRQERSKEKAKSRNENCLNRKTNDLFYS  
CERMMMLPQGLMASALLDIMRIPVLKTTGSKCMYLSNASFTEAEDDIVRLTSCLLNLET  
PGKHFSLLEKRRKILEYDSYNSMGNRKESKRWEDLLNVLKQHTNDENQTLNMFSDSDV  
LVKWNLMVGHHKNVCRLTGTQFKDSETFLKIGHVKFFRCMNSNSSGENQANELGGFAAKR  
RTKPNTIYNLAESPLMLSPESTLLIMLTGSDYNSAIVSNCEYDTWVRKEVAVFENTYCT  
CVGGWEIFLSEQEARKNNKDCDDSVGNISMGNLSKSNCRKCDKKLVLPFWTIKFFYLSQA  
IDFVRDPLQLCFPPTHLIDLETDVSLKHALHRAVNAANVMSYLTMGSFNQRFVGTIT  
TLSDISIHLSGANNNESKNTGSDVESDTEDLIPFSNNKRKSGNDPQKSTRKSKVNATRK  
SAPVTKKLSSSVFESIRGFFESHTEGGIINDRGILTKEIDVFGNNLDTNPEALGEENG  
GGGIVSSIPGLSTEQTSILKTEQNNSTSDFLDFFKKFNEMDDVEEEEEKMEEGEKEEEEA  
DLETDDWLDEARKAFEYKDSDFLEAVTAATNEMTSSLAKNNIEDEHSRCSVSSKLNNKQ  
PVMDEEKWAEIVNEFDKCISLDNITYNDNSLLSRLSGVLMANKREDGNNSNVVLYEPVQ  
GIDDERFSGVPYSVKTMNLLIVYMNMCGLDNTIVYQQLMPIIHSEFCGKTEEDKICTD  
RTNFMSAALEYTMLQYMPCLKKTPRIKQIKRKNWERIPKVLDDFKDKVSTCTDNYNKLLA  
TLNKEGKIPSENTKWLPSQGFMPVLGVAISKWSPPLTWSSFYLQHQQRDVSLTNITP  
PNSPRPEQ  
(SEQ ID NO: 189)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q12532 (Q12532) HYPOTHETICAL 119.1 KD PROTEIN YPL009C	50	1e-04
Q9XZI5 (Q9XZI5) RANGAP	47	8e-04
Q13387 (Q13387) HYPOTHETICAL PROTEIN 384D8_2	46	0.001
Q9YTL7 (Q9YTL7) ORF 48	45	0.003
Q25662 (Q25662) REPEAT ORGANELLAR PROTEIN	45	0.003
O96266 (O96266) HYPOTHETICAL 283.6 KD PROTEIN	45	0.004

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 860 of CT607: this corresponds to nucleotides 65102 to 65960 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=1428  
5'stop=1448  
3'start=1516  
3'stop=1538  
5'primer=CGATTCTGATGTGCTGGTGAA (residues 1428 to 1448 of SEQ ID NO:188)  
Tm5=58.31  
3'primer=TGGCCAATCTTTAAGAACGTTTC (residues 1516 to 1538 of SEQ ID NO:188)  
Tm3=58.35  
probel=TGGAACTGATGGTTGGA (residues 1450 to 1467 of SEQ ID NO:188)  
probelstart=1450  
probelstop=1467  
direction1=Forward  
Tm1=69.04  
score1=1.95  
length=111



CT609

Nucleotide

Genomic coordinates:

Start: 78365

Stop: 77441 (SEQ ID NO: 192)

Amino Acid

MWCSTHLSYSEFFTPLQSQKLFRNFFRALEFRGWTASSTECQVPRVDLWVGPMDSYTRNC  
WFQKRTLTFVCFWNRRFWRLVDPEMRGYNLLFSLENFTLPLSQKLKFNFFRALQFRGWTA  
SSTECQVPRVDRWVGPMDSYTRNVIAPETYINFCVFLEQAFLETGRPRNERVYPSVFTRE  
FYSSSISKTFQKFFRALQFRGWTAASSTECQVPRVDLWVGPMDSYTRNVIAPETIEEVSYGH  
FWTRCFWTKILLDGNPLPLPPPFKKGPRVYNDCTTPHSNHHNHHHHHGRTSILQQTLSRK  
WSSLL

(SEQ ID NO: 193)

Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
DYRK_RAT (Q63470) DUAL-SPECIFICITY TYROSINE-(Y)-PHOSPHORYLA	37	0.22
DYRK_MOUSE (Q61214) DUAL-SPECIFICITY TYROSINE-(Y)-PHOSPHORY	37	0.22
O61543 (O61543) BHLH-PAS TRANSCRIPTION FACTOR SPINELESS	37	0.22
DYRK_HUMAN (Q13627) DUAL-SPECIFICITY TYROSINE-(Y)-PHOSPHORY	37	0.22
HXA1_MOUSE (P09022) HOMEBOX PROTEIN HOX-A1 (HOX-1.6) (HOME	36	0.29
O60275 (O60275) KIAA0522 PROTEIN (FRAGMENT)	36	0.50

Comments:

CT1010  
Nucleotide  
Genomic coordinates:  
Start: 41348  
Stop: 41795 (SEQ ID NO: 194)

Amino Acid  
MTHLVLLILSLSPVYHHLTPYLSPHLTYTPISPITSIFPHLIHSLQFQHPVLAEPTHN  
QIWTPVFPFIPNRHHLCPQALAVYIRRRGQARSISLQASRRATQQALSLLLPRRDLPII  
KLQEWPLQPPPHQVLTPCWTLSTYLVLRN  
(SEQ ID NO: 195)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O62081 (O62081) C31A11.6 PROTEIN	36	0.20
Q05330 (Q05330) PUTATIVE ORF	32	2.2
Q60974 (Q60974) RETINOID X RECEPTOR INTERACTING PROTEIN 13	32	2.9
P79230 (P79230) KAPPA CASEIN (FRAGMENT)	32	2.9
Q9WAL8 (Q9WAL8) POLYPROTEIN	31	3.8
AAD12852 (AAD12852) Y8A9A.2 PROTEIN	31	6.6

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=209  
5'stop=227  
3'start=280  
3'stop=300  
5'primer=TCCTAACCGGCACCATTT (residues 209 to 227 of SEQ ID NO:194)  
Tm5=58.83  
3'primer=GCTGGCTTGCAGAGAACTGAT (residues 280 to 300 of SEQ ID NO:194)  
Tm3=58.29  
probel=TATATAAGGCGGCGCGGC (residues 250 to 267 of SEQ ID NO:194)  
probelstart=250  
probelstop=267  
direction1=Forward  
Tm1=69.00  
score1=1.99  
length=92

CT1011  
 Nucleotide  
 Genomic coordinates:  
 Start: 41758  
 Stop: 42097 (SEQ ID NO: 196)

Amino Acid  
 MLDSKLLSSEELKELTSYVSTSSRRSDMKHLLHLFEEHEKIFQFIQGHKFSLYTLDFE  
 IFYVMLNILLVEVKNILSPIPLLFDRNLQPVRRLWMFHNGPASPERCSRSLG  
 (SEQ ID NO: 197)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O44068 (O44068) DNA POLYMERASE ALPHA (FRAGMENT)	31	2.3
Q49563 (Q49563) DNA, TRANSPOSON-LIKE ELEMENT ENCODING 3 ORF	31	3.0
BAA74885 (BAA74885) KIAA0862 PROTEIN	31	3.9
O44055 (O44055) DNA POLYMERASE ALPHA (FRAGMENT)	31	3.9
O76063 (O76063) LEUCINE-RICH REPEAT PROTEIN SHOC-2	31	3.9
YD89_METJA (Q58784) HYPOTHETICAL PROTEIN MJ1389	31	3.9

#### Comments:

TaqMan Primer/Probe Sets:  
 5'start=107  
 5'stop=127  
 3'start=193  
 3'stop=217  
 5'primer=TCGAGGAGCACGAGAAGATCT (residues 107 to 127 of SEQ ID NO:196)  
 Tm5=57.83  
 3'primer=CTTCAACCAACAAAATATTCAGCAT (residues 193 to 217 of SEQ ID NO:196)  
 Tm3=57.46  
 probel=ACAAGGTAAGCACAAAGTT (residues 138 to 155 of SEQ ID NO:196)  
 probelstart=138  
 probelstop=155  
 direction1=Forward  
 Tm1=69.01  
 score1=1.88  
 length=111

CT1012  
Nucleotide  
Genomic coordinates:  
Start: 42053  
Stop: 45491 (SEQ ID NO: 198)

## Amino Acid

MAPPHLNAAADLLDKVMSGPLSPEGAQTSSPAACVGAKVVKALVSFCOKTRFTTNIVMRE  
VKAMEFQGDDFNYSALCASPQRPVTERQMFALMKSEDEEMGVSANFSPVSDDVINPSSL  
PSGQEVDSSTSAQISGMFQNVWSLLEECGSGSNSNSPVSRTVLVCTLFIIQVFKFLVTK  
VSNVNVNLNQLFGHVVFVGSGLDVAPSNNSVVPSTVVNNNNKPKSTSNNSNNISNKRVGGSNNS  
GGGRSKKVTATAKNPFNNVDGDNHGMFAGAPVDVNLDDFVFPQVETLTSKSTIPKEEVNV  
DEDLSKMKRKTALTPLEIHTFNVFISEINPSKYDRSMFCGFLTAWDKFVEGDTAGVKRF  
RNYILTRSNYASAAARAVYEASIKGTVYYNDKSKFLFDNVNPDLDKSWGKNKGKKPRLPA  
NLMAFMGIDIVKCAKGIQKYMFAKQFQHPEVEELVPPMAVYAKVAAGLKSGTLFDDWDL  
PEYENCQFIKYDTEGCKKHSELYAKQLLRTGLNQYNKLEEGQSAFFANIVTVTSASSDD  
IHGDTIIELMYKTKDGVKGVSKIEDENIIKVNPAEEKNNRVQAEKTLTYFEIDSDEVCE  
RTEEEFFRPTSVVAAPTTPLVPSNVEEEEEEEEQMEEEEEEEVEREEGSDKEDDGDAPAQ  
EEMEEEEKEEEQQQQPEEESNGNENQEEEQQQQQPEREENKDAQSDSDSDSSSSSSSSSS  
SSSSSSSSSSSSSSSSSENEAEKKKEEEVPAKIQKRKRLSERPSEAASSPKMRVVEEQQ  
QQLSPSLDILQTAVDEMMEEI PAPEPIVATTSPKAATLALKTGFSYSSFVRGDDLSVAGN  
TSPTPEAAVPAATACTSDVGNDFLDMLDGLPGDIVMQPGCEDVTAKEFFEGITLPDGTNE  
CTGFDDLLKATETDNIITTTCTFSPIHPSSNSAPRKDIDNCSSIKRSRAGSLFDTDDSE  
TNEVEKEAPKRKKHLKRRNKSHRGSSGSASSSHCMSSDEESEDERDMKSTSKVHKSPKA  
HVKHSPKYDAVNSDVNNSYNNVNSTTCMSSSDSDAEAPKSHNKSRSRKHSSSSSTSDDKKQ  
NQCSINTQNVKKTVVQSPPSFRSFSPPKKDELGDFLSRKHTKPVRYNNKKRDVNTTNNV  
VQRSA  
(SEQ ID NO: 199)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q9YTL7 (Q9YTL7) ORF 48	101	4e-20
O95815 (O95815) DENTIN PHOSPHORYN (FRAGMENT)	99	1e-19
GAR2_SCHPO (P41891) GAR2 PROTEIN	97	6e-19
AAD46501 (AAD46501) LATENT NUCLEAR ANTIGEN	94	6e-18
O40947 (O40947) ORF 73	92	1e-17
GARP_PLAFF (P13816) GLUTAMIC ACID-RICH PROTEIN PRECURSOR	91	3e-17

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 955 of CT1012: this corresponds to nucleotides 44515 to 45469 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=1345  
5'stop=1366  
3'start=1389  
3'stop=1407  
5'primer=CATCCGGAAGTGAAGAAGTCTTG (residues 1345 to 1366 of SEQ ID NO:200)  
Tm5=59.57  
3'primer=CAATCCTGCGGCAACCTTT (residues 1389 to 1407 of SEQ ID NO:200)  
Tm3=59.65  
probel=TGCCTCCTATGCGTGTAT (residues 1367 to 1384 of SEQ ID NO:200)  
probelstart=1367  
probelstop=1384  
direction1=Reverse  
Tm1=68.99  
score1=1.99  
length=63

WO 01/38351

149/201

PCT/US00/28888

CT1013  
Nucleotide  
Genomic coordinates:  
Start: 45950  
Stop: 47825 (SEQ ID NO: 200)

Amino Acid  
MSSTTSPSSSLWDDDDDDDEDEKDVKQEVSNRPPIFSYMETVSFSDNDEDDNKGEEECFG  
SNFDMFGDSDNMPSTSTAPFPPPTTTPLPTPRSIMDTDSDECDEEGAAAASAPSTAAASS  
SIPVGISEAELKKMEKKKKRKEIKKLKMMKDPLPHLYVGGEPPVAADYKTRANISLYKVD  
PSIDMCGVAPPQFCAELPTPSIDVYTSSYVFPPTPAMHNKKGSKKCQFLKGRKALRKWI  
HENVCMAAPPGRGGVFLAHLERFLAHEGDEYKVRPMFVSRVLNKAFFNLARADTLCS  
MTFYTNLCWIVNGVVVCFDKDDGGIHDASEYATGENFDTVVFHKREEQKTNGSASKRR  
LTPDTSNMGSTSDVQEFQTMGTNTDMQEFQSMGTNTNPIETSSVGVNTNPLPNPPRLVI  
TPLTNDVPELDMMWLYSPSRGGGNSRMSANTGTSPLSNTPIPTCFTGGANVVVPNGFVPP  
TFPLECEDDDPSIPNSYNYEEDKVHFPFYEYMAKYLSPLVPSYNGQTCNVVQEWFKGSF  
SLAKRRGTVPKFCNSHAFECNMDVCTAMCKWAKTVIRHGQYCNRCIVRRSCTSMLAYH  
YIVCRDASCDVPKCRERVNRNDMDD  
(SEQ ID NO: 201)

Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
AAC97971 (AAC97971) FAS-BINDING PROTEIN	47	4e-04
Q19871 (Q19871) F28D1.6 PROTEIN	44	0.004
NUCL MESAU (P08199) NUCLEOLIN (PROTEIN C23)	43	0.006
AAD56625 (AAD56625) NUCLEOLIN-RELATED PROTEIN NRP	43	0.008
VP41_YEAST (P38959) VACUOLAR ASSEMBLY PROTEIN VPS41 (VACUOL	43	0.008
O35613 (O35613) FAS DEATH DOMAIN-ASSOCIATED PROTEIN (DAXX)	43	0.008

Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 1 to 934 of CT1013: this  
corresponds to nucleotides 46867 to 47800 of the genomic reference sequence.



CT1016  
Nucleotide  
Genomic coordinates:  
Start: 52093  
Stop: 54913 (SEQ ID NO: 202)

## Amino Acid

MSEPSVYAFIDIKEIENGWEKEFGLLVQPGQKLAPFRDISYDSSKLDCAFSCIPSDILH  
SDNEKRVGECNFAHTSVSFPVKNPEGKTLRHFTACGPGCYRRYKORDPHTGLPVARGVL  
MQDHVDHETGNKMCEYLNQSLVMWAAPVWIRPGDLTEGYNTTHVPGFAFKEDDERDSKRV  
KYENVVISKAYCDEFFKQYYDADSGSCYRSGWMKFVHLMFGQYFTNLSYNLANPKPYNLTG  
NTWSDVVSVLTDNPIVDAGAAPSRSSEMDEIITKKKFNVFPSEQTSARQKAENIIRSQYGD  
GVEIDPSSVDALMQFVNREGVVGTEKKSDRLMRVADAVMDAAMRLQVMGLDDSQSRLLL  
KNMIKMSRNNPEYARHFSSSLKLIGVTLAIKRSVFSKGASAKRKETAINNGEQHRRSRWS  
PETVTEEDALLFARENITEDPKHPAPFVDILHSPDINSSIKSGSSSIWNDILSRISSTR  
KLEEKASVFKNLVVKVVRQFLDILEGKLFS DGYEWDDNIPLMIGVDQILREVIKAASN  
CARFASSALESSLVTGFIDSASAITSR LAVQLAARTFSVFLEESVIEFVVAASLR LAIQA  
FADLATLAASALTIVIGIVIFVIOVLGLILDALGLGWYDHIFSPEDLKKQVLVFRREFAK  
AGNV DVGVAQPVTPPEEIVAINVFLQTEENGEEKKEEGARKSKIDFLQKYFHSTPLMGKKS  
KFVYIQEAAQEYLGGR TMNAFGQRIITAADDSDTTTTQEGRRDDET VTKMRSIILETG  
QTLKDYSSAVNYNASRLDYVGEEWVRNTALKEETRSNTSDNLFKKT VSLAS MAGAFLVL  
GIGVLVASHITLLRFTNIGLAF AFAGLLAFIALMSISYINMNAMGVVNSDAIYRSTALVG  
DIKTDPRRVGMVQRHVGVGAKYNMITDFVSPMLDEIESD  
(SEQ ID NO: 203)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O67140 (O67140) L-SERYL-TRNA(SER) SELENIUM TRANSFERASE	41	0.048
YD86_SCHPO (Q10411) HYPOTHETICAL 222.8 KD PROTEIN C1F3.06C	38	0.32
Q88606 (Q88606) (CRUCIFER) GENOMIC RNA FOR RNA-DEPENDENT RN	35	2.1
PME3_PHAVU (Q43111) PECTINESTERASE 3 PRECURSOR (EC 3.1.1.11	34	3.6
O81424 (O81424) T2H3.2 PROTEIN	34	4.7
O04260 (O04260) HYPOTHETICAL 49.5 KD PROTEIN	34	4.7

Comments:

CT1017  
Nucleotide  
Genomic coordinates:  
Start: 55054  
Stop: 58189(SEQ ID NO: 204)

Amino Acid  
MDSLNTNTVTLVNDRLGNHRTNKPITEQDVENTLNLNSLERASLLKLYSVFIKEMQSYSG  
CIPKNKYTNVQEIFEDGLITFEWRDGTKVHRSVSPSSPIPLSTKKSPPSSPSPSPSPSI  
KEEEFEEEFEDDEEIIYETDENVEDFINGDGEDSEEEEEEDIIVDDEEEENEENKYNVLA  
FSNHLRRQTAAAAAADIEKKDKNHAVSAHDYTLALQQQQQKLLQQQQQQHQ  
QRSSEKVTSTPNKFNKFLPSNGFSEQTELFVCFDVKIAQYNGLVLDILPIVAEYII  
NGLGLKCSMETPPVKPCRRKEVKDVWCQPKTSFENDAVEDKHLAFAESPILQRPRDFPI  
KKITAYFCLDDSDVIKNPWGSCPLLKSGSNFRVSEYSRHFNEFSGVKNDDDTSSNTCFIY  
SQKNPNIEIVSKLNIEFEVMEGIITHRKDLFETGILSDSSLATAMAFCHPKARVRNVAL  
FYFSVYLPFSKITRKETIKCSETDKVHIGSDAIFSPPSDNPNIASHQNNNNNNNNNTSVN  
IEDRPIRNNNISRKMTITNYQCMACKERCTNNCTNGNYPDRGNHLSHSHVKGEDFFKILN  
NSKVDLSLKKLSRVLIAPPSSGNYTSKFCDRSSMCHSFFCRGIEPVSTSFSSDSFEKTKLV  
LYGKVVDVINSYSAIKTSHNNRIRVFFNSEEKDNKTIPSRASAKNAFKDILVHECNKER  
AVSYFEQNKLSKDGHLNKNWWIELNDLNMFEKHVEDFYKKCSKVND AESLKDIFNDFE  
KTCDKYKTAKRAIIGAQDPSTSTPSKKENGITRIISTLSEFHSKDEATVSALLDKTMLLG  
SRTIMSGVRCVIRNNNSVFSGFENKNTNNNWELEIRHYVISMGGAATKISDEDELEQFTPV  
RGAVSVTTAPNDKLPVGAHQTKWDEQTLKTNTKRNSLYDSYNKRNNRDNNKIKNRSKL  
SDFNWRTPNISIQEFNANKDDVNKKRYAEVVASAAPKSPSPSTSSSSSNSNSSPPLSPLS  
PTVKNSNNKPLYIPPHKRMTTAV  
(SEQ ID NO: 205)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
077033 (077033) TRFA	60	7e-08
Q94463 (Q94463) K7 KINESIN-LIKE PROTEIN	59	2e-07
NUCL CHICK (P15771) NUCLEOLIN (PROTEIN C23)	57	5e-07
AAD46501 (AAD46501) LATENT NUCLEAR ANTIGEN	57	8e-07
O40947 (O40947) ORF 73	56	1e-06
Q98148 (Q98148) ORF73 HOMOLOG	55	2e-06

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 1 to 215 of CT1017: this corresponds to nucleotides 57667 to 57881 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=1741  
5'stop=1762  
3'start=1847  
3'stop=1864  
5'primer=CGTGGTAACCACTTGTAC (residues 1741 to 1762 of SEQ ID NO:204)  
Tm5=59.16  
3'primer=TTCCAGAGGGAGGACGG (residues 1847 to 1864 of SEQ ID NO:204)  
Tm3=59.84  
probel=TGAGCAGAGTACTGATTC (residues 1829 to 1846 of SEQ ID NO:204)  
probelstart=1829  
probelstop=1846  
direction=Forward  
Tm1=69.01  
score1=1.98  
length=124

CT1018  
Nucleotide  
Genomic coordinates:  
Start: 58947  
Stop: 60029 (SEQ ID NO: 206)

Amino Acid  
MKICQISSPTLTLSIPLEGVYHVKQLLHLKVHLDVKGVKQLLHLKVRLDVRGAKQNPWRK  
NLCLLKKNVKSQKQLPHLKVHLDVKSQKQLPHLKVHLDVRGAKQLPHLKVRLDVKSQKQL  
PHLKVHLDVRGAKQLPHLKVRLDVRGAKQNPWRKNLCLLKKNVKSQKQLPHLKVHLDVKG  
VKQLLHLKVRLDVRGAKQLPHLKVHLDVRGAKQNPWRKNLCLLKKNVKSQKQLPHLKVLL  
DVRGAKQLPHLKVLLDVRGAKQLPHLKVLLDVRGAKQNPWRKNLCLLKKNVKSQKQLPHL  
KVLLDVRGAKQLPHLKVHLDVRGAKQQQQLCLPLKTISTSFTHLLCLYMEYGKHQNLQV  
X  
(SEQ ID NO: 207)

Top Blast Hits

Sequences producing significant alignments:			Score (bits)	E Value
041125	(041125)	A643R PROTEIN	36	0.34
077336	(077336)	PFC0425W PROTEIN	35	0.75
Q65683	(Q65683)	42K TRANSPORT PROTEIN	32	6.5
Q48275	(Q48275)	HYPOTHETICAL PROTEIN (FRAGMENT)	32	6.5
Q89659	(Q89659)	42K TRANSPORT PROTEIN	32	6.5

Comments:

CT1019  
Nucleotide  
Genomic coordinates:  
Start: 62203  
Stop: 63019 (SEQ ID NO: 208)

Amino Acid  
MDVEFGFFHGLLSKALLPDEKHQPVIRRLCADDNRKGEDGCCSFCGRRGTGESNTACLE  
QLIDVCSFIGTVSSIGTIINSNLSTSCSRLQKTSDSAALSHSSFLDVVYPSLKKTTEDV  
LPHSLRAIWNKQLPKLYEKTLPPIEEEDIGYKDYVVSIEDDDNVDDGDQEQMIIDEESY  
KTIGEKSTIELIGMYNNKFGNEFIRIPLRETALHAQSLRYDTEAKFVNHKDSIPLFYEN  
STCTCKERLIDFSERQLQQLKQDGMKPTDK  
(SEQ ID NO: 209)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
AAF02799 (AAF02799) F5I10.23 PROTEIN	37	0.14
Q23058 (Q23058) BAC IG005I10	37	0.14
OPI1 YEAST (P21957) NEGATIVE REGULATOR OF PHOSPHOLIPID BIOS	34	0.90
Q24042 (Q24042) HYPOTHETICAL 89.2 KD PROTEIN	34	1.2
Q24043 (Q24043) HYPOTHETICAL PROTEIN (FRAGMENT)	34	1.2
O76871 (O76871) EG:100G7.2 PROTEIN	34	1.2

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=292  
5'stop=313  
3'start=375  
3'stop=396  
5'primer=GCGGCATTATCCCATCTAGTT (residues 292 to 313 of SEQ ID NO:208)  
Tm5=58.17  
3'primer=TTGTTTATTCCAAATGGCACGT (residues 375 to 396 of SEQ ID NO:208)  
Tm3=58.03  
probel=CTGAAGACGTATTGCCTC (residues 350 to 367 of SEQ ID NO:208)  
probelstart=350  
probelstop=367  
direction1=Forward  
Tm1=68.97  
score1=1.97  
length=105

CT10000  
Nucleotide  
Genomic coordinates:  
Start: 3143  
Stop: 6956 (SEQ ID NO: 210)

## Amino Acid

MVLTLSCTTRRVASSKGNFSKEDAVLGNQFPILKKSNNLSIARPPSIESFSASVEKIFRE  
WNESSGGEKIFDISQNEEEWMDIISLVESVYEPVFSKSLKPKDLADKTCLTAAAFALASA  
VDEKLTILSGSDGSLVLTQRTTKVMKKDPKKIAESLLNNEKWTSSILLDRKLTAKKLLSRRGA  
LKSAERVEVLHRLNKLKEAPLPHHPSLFDNFSGGKTSAVSAGTVIASDMHFKLVEHIFKV  
SFRKWGPCGDKTESGEEDEEEEEEEKKHSISRFLVQFMNGHNGQHYHRPESASVYFCDY  
YDYLAYRNLNNEYKLSSMHPGTENMEDLPFRPFVAVPSTYKTELEYKRFVQSTNLPQLSFD  
YGEFLCYCIFGADWYKHLGDVVDLSLENSSMISFDSQTLSGVYKNTANYKRLGKKRNGIAD  
LAVRSMAEFIRTEAHKALTAEMEAAAAAAAAAAAAEAMDQEPAEVDFLSVPHLRRKIRQAV  
SVLNNFVENDLSILVSNFKNVLTDDTVSGTDTDNFGSSGEFEALSSHFLSRILDEVHIL  
RNTDIQRTLFSTHVSLSDKSPPSRVRGSSNVNFNNNAGNISSLOQTYGGIEELPENVLVGLS  
GGFEDTDMYSGEDVVVVWDGCDGGKVLSTFNCNDFIQLHEKTAETFKDDTDLVERIRD  
VLQTASKTGNLKKAYSARKNIYAVLRENGIERPGDDFTEKGIALKDKTNQPPPPPARSAKI  
TVEGVKGGFFSGFRDILETRALTYSATFRDLGQIVKETEGTLTAATVAETSFSSEGLAES  
LRSDANLGLFESEDAKTVVFNKNTSRSLLEETRALRANNTSFSSSFARMDGMGVQSADLDAE  
FAEMRETYPDAALEQNLKDLKFEETIPESQVKKLKKIDSYLTHENPERAGKEINDTELS  
KATDSVLGKKLGNVAVTLMNNGKVTIVVGASVVAGFLGPAVALVHASRGHNLNVVDHT  
SPKGVISYKIVDFSCADRNTGWAKPTKHPFREEIDHVIALDASFLTENGAYVFPEDGGPK  
SKYKAYAPICGTDAAQGECSWATFDDPHSVLPWVASMKDLPGQSLSCDKGMSTLKAV  
SSVLLSIGKDVAAEAFIEVAEDAVVGLASKAISAVINNPLFIFGVPLGFGIAATRLNPSNW  
KTGLIVFSILLVILIVRFFAGSGPLTLNWFGAKNSAKRKQTEQFEDGGGNRSKIVLAEK  
DNANSKLQSRNETGPMRLEELPGHEDLRPVFFPATNYSKSAKILGYKSKPFDNFYTKI  
INTDIIKMDR  
(SEQ ID NO: 211)

## Top Blast Hits

Sequences producing significant alignments:

Score (bits)	E Value
-----------------	------------

YK06_YEAST (P36062) HYPOTHETICAL 75.5 KD PROTEIN IN SDH1-CI	43	0.010
Q12532 (Q12532) HYPOTHETICAL 119.1 KD PROTEIN YPL009C	43	0.017
O75184 (O75184) KIAA0702 PROTEIN	41	0.050
O60721 (O60721) RETINAL ROD NA-CA+K EXCHANGER SPLICE VARIAN	41	0.050
O43485 (O43485) RETINAL ROD NA+/CA+, K+ EXCHANGER	41	0.050
AAD28522 (AAD28522) FLAGELLIN (FRAGMENT)	41	0.050

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 969 of CT10000: this corresponds to nucleotides 5867 to 6835 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=2102  
5'stop=2125  
3'start=2185  
3'stop=2204  
5'primer=GGATTGCTCTCAAGGATAAAACAA (residues 2102 to 2115 of SEQ ID NO:210)  
Tm5=57.56  
3'primer=ATGTCACGAAAACCGCTGAA (residues 2185 to 2204 of SEQ ID NO:210)  
Tm3=57.81  
probel=AAGTGCCAAGATAACGGT (residues 2148 to 2165 of SEQ ID NO:210)  
probelstart=2148  
probelstop=2165  
direction1=Forward  
Tm1=68.99  
score1=1.99  
length=103



CT10002  
Nucleotide  
Genomic coordinates:  
Start: 20783  
Stop: 23729 (SEQ ID NO: 214)

## Amino Acid

MEETITLKESTGTVSPFKNENITRIASNYVRAFTDTWSHLVNISGAPLTAEKNPSAIPAN  
ELNRYWTKTNVLCNPLFKLEDHITRDEDTGTITLKFKMYIDDKNGLYQSAVLMALDSFV  
SLASFHSGADLVSNKSENKFCVKIPHDTRAESLLNNVGFPAGLSGPFKRWSINYKAANLS  
GKSGIDGLSGSMLTVLKNNTNKRATDILHLVNNVSASAAQQLDDSEMSRTFNHQKKVVCY  
DINVSSSRQVNQRNLLHHQNIIGQHLIEFRTKQLERAQNKVKKEEENGEHEMTSEEEEE  
EDEYEEGGCLSDIDEEDFYEDGYDEEEGDDNRTRKKKKMEDEDEEEYDDEDEDEEAE  
TCGANGVIDCEDDAIIFPNGQNSKRKKNKGKKTNIKRSRRKGECSANTLSFVEKYVGNCK  
SLGIKPVGCPPPSTEFTSLFMKGSEADSCYNTCQSTRGASRIRSLLNKYSVKDLMQVNSP  
SSWKWANPPDRRFVLFDKKTKEEVEVKFEIECEKSEYFDVVSELPSNIKVWLKETAKIHK  
HLALIEDFLPAMGAATPKIPLNLIKTMTSIFSVRDIVGFKIPEEVLSFIPIEWKTSISAM  
GLLSVQFDRIIEVIDLMITNGAFATSCLNNAFFLERGVVPRDGSNTWLHTDLVQLSTSIF  
RSIRNRGVNIGGNNTGSNSSSSSCGGNGGDYGVRCGLSISKRGITLKPPPAAMTNSSSP  
SSSAMISLPQPTRQSIDLSITITIIQDFSEVSGKLRNLGLQKNMSDKSKDVFNDAIYDSGA  
FKALLTCTVNDKSRKRKRRTLLASGEGVVRNLMVMSQGNVDNDAHQFQEECGIKIGGGA  
SRVYKRAQRSGSAVSSRRRVNRKPOFTIAVSDEDDCEEGDFSSELNPTHSQLLLFQQR  
QQDSCTEDDDLVLSVEEYNNRVSGSSTTAGDRVLAkdLLSTVSPNEKRNSAALAALTISR  
HSLFNALSAKTKLGENGRFFL  
(SEQ ID NO: 215)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q9YTL7 (Q9YTL7) ORF 48	75	2e-12
GARP_PLAFF (P13816) GLUTAMIC ACID-RICH PROTEIN PRECURSOR	75	3e-12
Q07034 (Q07034) RNA BINDING PROTEIN	74	4e-12
NAB3_YEAST (P38996) NUCLEAR POLYADENYLATED RNA-BINDING PROT	74	4e-12
Q9YPA9 (Q9YPA9) HYPOTHETICAL 45.2 KD PROTEIN	72	2e-11
O40947 (O40947) ORF 73	69	1e-10

## Comments:

## TaqMan Primer/Probe Sets:

5'start=1699  
5'stop=1723  
3'start=1773  
3'stop=1797  
5'primer=ATGACGAGCATTTTCTCTGTTAGAG (residues 1699 to 1723 of SEQ ID NO:214)  
Tm5=57.07  
3'primer=TGCAGAAATAGATGTCTTCCATTCT (residues 1773 to 1797 of SEQ ID NO:214)  
Tm3=57.23  
probel=ACCAGAAGAAGTGCTCAG (residues 1743 to 1760 of SEQ ID NO:214)  
probelstart=1743  
probelstop=1760  
direction1=Reverse  
Tm1=68.95  
score1=1.95  
length=99

CT610  
Nucleotide  
Genomic coordinates:  
Start: 85707  
Stop: 83427 (SEQ ID NO: 216)

## Amino Acid

MPPKHKPNTALKKHIIRNQQRKKEDDAESRFQRNMGOEVSKLDAPTSSKNRQRRKIRTSK  
ILSRSGDCVAGDCSDLNDEGKRDTDOEGGGRGGGNEEEEGKEEGEGEEQQREEKEEQS  
EEKEEKDGEEEEENVEDEHVTPTTSVSKRAKQMKKHIFPPSKKRKRSDESALAVPAG  
KMMTVSRPLRGAITSGSILGVRSENAPQYDYVSYLADEAVVKEKAIQYRIRSLANLLKA  
NKTKAFPTSSSLLSSEQGKKKFGGKRTNTFVVTNVGAELVKALLANSCWAISHRKDIRSG  
EIQWQELSSKILKSLNDGNATEINNLMSSIVEDRIQRTVKERVYFEQLATVCNNLFGTRI  
LPNKNFDKNFVSVASDNSNATVRGLSIPRYFRAINNNVWVKMSSTMDLLVGGGMRSKSEH  
SISMLEKCAAGVLARASARPVEKMIKSAVEETSQAFNLSTGVFVPKQQQQRQQQQQQFP  
PFQPPPPFLPPPQAPFQVQQPTYQGYLNPYYQYNQYNNPYAPQQLQQQYPLYFLGNQSQP  
PPQLQQQQQQQQFPQPPNNIPPPPTPQQQSPSNIPPPQQQQQQFPFVQLISSPPPPPIIP  
NTAPSPPISRVRFDSRSTTPQPPPTPVLPKPTPLPPPSTARAEENATDMSFTDIDSELG  
SIDFDLPPATPGRNVEEIIKAQRQAVKETGVRGEEEEEEEAFAPIIRQPRTPGNFRDEL  
LDVNESIYGSDIEPAAAAAFDWDMLDDDLNGDEPYEFE  
(SEQ ID NO: 217)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
P93794 (P93794) LOW-MOLECULAR-WEIGHT GLUTENIN STORAGE PROTE	84	3e-15
Q05573 (Q05573) OMEGA SECALIN PRECURSOR	83	5e-15
AAD43602 (AAD43602) T3P18.1	83	8e-15
O04365 (O04365) OMEGA SECALIN	83	8e-15
O48809 (O48809) F2401.18	83	8e-15
O65375 (O65375) F12F1.9 PROTEIN	82	1e-14

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 34 to 361 of CT610: this corresponds to nucleotides 83630 to 83957 of the genomic reference sequence.



CT611  
 Nucleotide  
 Genomic coordinates:  
 Start: 88938  
 Stop: 85761 (SEQ ID NO: 218)

## Amino Acid

MEGVTSIVAAVPEVAILITDLMGGRNNKRSTYERIVGIVGESGDLLEAILDICNRNSYR  
 DELLEGETVINPTGLLKEISLLMKKALDMNIKMSNDPVPFTTLDQNEQEFIGHLKSC  
 KKQDGPAYKDLIHRIYSGMFVMKNTRLMLDEIIRGNAGDAVEEKNALCEAYAEMISDMDL  
 IRIFLLLVAIKRDQNKKRRHMKSVIYEDVVVSLNTLKDVFHKEWYMWPFSAALQVGTKIRD  
 ARTFSVLFGSDMHEGRNNDRIWENMAFSVTEAFLSGPSTNNHYNKGHLRMYAARPVYDAM  
 EYVPQELHHILFGTKIAKMIDIVYRSIYNVPYLLAADTERVEEPKKSVMSPSGLIISPN  
 ASLLENTPLSLVSRHGIPSAKLGFILEHENAENMHLEEAIAKCMVSQTLQEESWGESQ  
 AAMVYQPSDEVEVIQAHVTKILSGNTTNKTCGLCYADLDMKPKFFNCSENMKASYDYFP  
 VHAFMDTFEARQETCSAKLCPDCTIKHLMYVYEKVSAGSEKLDVFRCPCCGEYMVQFIG  
 RCHEFSSSLFERAILAGENVDPYIAANKLLITELIKRAEKCFTVELLQAEFMECKMDK  
 DFALDKDSKFTVVDNRFRPPVKLFKMVEGETGDSKCSLICTQCLLPNVCDQPNEMEDIVT  
 VDVPVPLVPYPPEQLEDYFQDVEDAEFDDPPTDELVRDYDTGPGHLKWPMLSCGFLAS  
 NFVPPNEEVTNCRQAVSILKRTPEKKIRGWNPESEPGKVLLALANWHSTDRMPENMKGLL  
 NDISVIHNTRERFQNRVKVHYLNSVFGGFDDRDFFEQVVGVSIPLIATYFYVYEKLNHESA  
 LGLWAKMFVKNLIGEMVLERPECVFHRAHSEFVLHCVDRRALSGIRENQAKMEIVKQVNI  
 VRQNMTSESIDPQVFTVDEKRTLEWKVEKEGQEIKTVKCPKCKTPNIKLGGCITMTCYDC  
 SGRRDGYPTVFCWICEDEITNPDHILIDHKLLYSDCKSTKAALKVYNCTLCCLALRKCS  
 DSYLSKQRGGGEEEEIEIYVMEDGFEF'DVHTKTAVPTK  
 (SEQ ID NO: 219)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
AAF04637 (AAF04637) HYPOTHETICAL 84.4 KD PROTEIN	1495	0.0
BAA78677 (BAA78677) HFB30 PROTEIN	44	0.006
AAD21842 (AAD21842) ANDROGEN RECEPTOR ASSOCIATED PROTEIN 54	44	0.006
O94793 (O94793) HRIHFB2038 PROTEIN (FRAGMENT)	44	0.006
Q9XII0 (Q9XII0) F7H1.11 PROTEIN	41	0.042
YK27_YEAST (P36113) HYPOTHETICAL 63.6 KD PROTEIN IN YPT52-G	41	0.056

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 829 of CT611: this corresponds to nucleotides 85988 to 86816 of the genomic reference sequence.

## Hit to public SBV sequence:

gi16165655|gb|AF099142.1: CT nucleotides 1 to 2738 match nucleotides 8582 to 11319 of the public sequence with a 99% homology, a score of 5398 and an Evalue of 0.0

CT612  
Nucleotide  
Genomic coordinates:  
Start: 91607  
Stop: 89060 (SEQ ID NO: 220)

Amino Acid  
MGSNQQQSFISKRNQKQEIISLEKIIKRIENACLPVNQYVPKLDKNAINPQELASHIMDR  
LPATISFQEMDDFLADYAKTKIVDHPDFGKLAGRFICSNIHKNTKEWNSFSATTQKLRHA  
IHPGTGKPASVVNDTYENVMANAEILDVIDYKMDYLFTCFGLRTLEYSYLIKIGSPTD  
RKKRILVERPQDMIMRVAVGIHGS DIKSVIETYDLSRHYFTHASPTLFNCGTVTPQLSS  
CFLGLQDDSIIEGIYDTLKEAAIISKTAGGLGIHFHDLRAKGSPISSWSGTHPGLMAFLQ  
IFNVSVKKVSQGGDKRRGAAAIYISDWHLVDVKDFIDCRKNAGNEDLRTDLFP AIWVSDI  
FMERVKAGKNWSLMCPHECPGLSDVHGEEFKALYEKYEAEKGKKEVVKARALFDQINSAR  
IETGTPYVCFKDTINRKSQENVGIIKSSNLCTEIVQYSDSEETAVCNLASIAVNKFVKY  
SPIPSLRPYVDYREMKRVVKIMTRNLDKVIDVNFYAVDKTRISNMKTRPMGLGVQGLADL  
FFKLRIFFESEEAALINKRIFETIYYGALEASCEIAKEKGETYELFEGSPLSKGIFQFDM  
GKENIKNRDIYFNSLPIHDWEQLRRDIMKYGVHNSMFVAPMPTASTAQILGNSESFEPLT  
SNMYNRNVLSGSFQVVNEYVIRELIKLGWNSVTKQIRIMASGGSIQTLNIPKSTKELFK  
TVWEINPRTLDMAIQRGMEVDQAQSLNLFVEEPELSKVRSMYAWEKGIKTLYYLR TK  
GAARAVQFTVDKNVLQEVKKEAPSPVAAFSAPVREEEEEKKSSIVVPDPAALLCSINN P  
GACEMCSS  
(SEQ ID NO: 221)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
AAF04636 (AAF04636) LARGE SUBUNIT OF RIBONUCLEOTIDE REDUCTA	1720	0.0
RIR1_SCHPO (P36602) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LA	807	0.0
RIR1_MOUSE (P07742) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M1	804	0.0
RIR1_HUMAN (P23921) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M1	803	0.0
AAD37491 (AAD37491) RIBONUCLEOTIDE REDUCTASE M1 SUBUNIT	801	0.0
RIR1_YEAST (P21524) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LA	796	0.0

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 845 of CT612: this corresponds to nucleotides 89225 to 90069 of the genomic reference sequence.

#### Hit to public SBV sequence:

gi16165655|gb|AF099142.1: CT nucleotides 1 to 2547 match nucleotides 5913 to 8459 of the public sequence with a 100% homology, a score of 5049 and an Evalue of 0.0

#### TaqMan Primer/Probe Sets:

5'start=1327  
5'stop=1348  
3'start=1383  
3'stop=1402  
5'primer=GTCCGCATCATCAAGTCTTCAA (residues 1327 to 1348 of SEQ ID NO:220)  
Tm5=58.78  
3'primer=TGCACACTGCAGTTTCCTCC (residues 1383 to 1402 of SEQ ID NO:220)  
Tm3=58.45  
probel=TGTCAGTACAGTGATTC (residues 1365 to 1382 of SEQ ID NO:220)  
probelstart=1365  
probelstop=1382  
direction=Forward  
Tm1=69.00  
score1=1.99  
length=76

CT613  
Nucleotide  
Genomic coordinates:  
Start: 94397  
Stop: 94175 (SEQ ID NO: 222)

Amino Acid  
MVIRLCFLESITCFVYGIMAPLSLDTNTDYLSHKKDTNKKIQMQINFIPYSNMHVYIAGV  
YTFHEKKGLTYQQY  
(SEQ ID NO: 223)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O96794 (O96794) AMINOPEPTIDASE P	29	4.1
YA2G_SCHPO (Q09706) HYPOTHETICAL 157.7 KD PROTEIN C2F7.16C	29	4.1
O24567 (O24567) ESR3G2	29	5.3
AAC36183 (AAC36183) PUTATIVE PEROXIDASE	28	9.1

#### Comments:

Hit to public SBV sequence:  
gi|6165655|gb|AF099142.1: CT nucleotides 1 to 222 match nucleotides 3447 to 3668  
of the public sequence with a 93% homology, a score of 357 and an Eval of 1e-  
101

#### TaqMan Primer/Probe Sets:

5'start=48  
5'stop=65  
3'start=169  
3'stop=194  
5'primer=TGGCATCATGGCACCCT (residues 48 to 65 of SEQ ID NO:222)  
Tm5=57.92  
3'primer=TCATGAAATGTGTACACACCTGCTAT (residues 169 to 194 of SEQ ID NO:222)  
Tm3=58.55  
probel=CTTTGGACACAAATACCGA (residues 68 to 86 of SEQ ID NO:222)  
probelstart=68  
probelstop=86  
direction1=Reverse  
Tm1=68.82  
score1=1.22  
length=147

CT614  
Nucleotide  
Genomic coordinates:  
Start: 129006  
Stop: 127158 (SEQ ID NO: 224)

Amino Acid  
MEYIGKNNNPVSNESVSEKELKLRSSFLMIGKKTSKYEQVMGVYEAIESIRQSELT  
FVVHVKKDKQLKFARGLKRLQELVEDDSLRIERISCAPPEPGHLFKDDAGHVTDEEWLAT  
QEEDVRKINTIVKEKLKRKDKDFKFSQLYRYMSNSLSEAVEKKHDCMISSDFLIGLGFS  
TMNVTHALKSMERTMQKHGFKDMMVPLVEICRTHYKGEYIANPIFKSHSSHCLIVPLFM  
VAGVFARSAHPSAASIEMYLSLAYAVILYSDEKQIQIREELARKNLQIKEELENQVEKT  
TKVEKELETQVVKTTKVEKELETQVVKKEEYKNSYIETEQLFKVSEEQKESLRNVHKKSS  
NATFRYDSGSLVFSISSTEFYLLCRTDKSGSFETATENGLRYIFSPINKKRDITAGMRPR  
LIMAVTGCDAPACNDISIKHQNKFVKLKNRSSIVFQTPPSDEDLKGIVQKVTGSDIRIF  
MNDGTVYQDGGRIDISSPQELDEENMTQFEIEQQRKLHSMMENTS KIVTRYNKERHLTTK  
EARTRNKTEKWFEKVKKREEQKKRENGEQSTSEQEQRGVKRTWENDNEFDSDVEEEDGN  
NTQEQQRVKRHAISV  
(SEQ ID NO: 225)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q06166 (Q06166) MATURE PARASITE-INFECTED ERYTHROCYTE SURFAC	54	2e-06
O66878 (O66878) CHROMOSOME ASSEMBLY PROTEIN HOMOLOG	53	4e-06
GLE1_YEAST (Q12315) RNA EXPORT FACTOR GLE1	49	8e-05
O35788 (O35788) CYCLIC NUCLEOTIDE-GATED CHANNEL BETA SUBUNI	48	2e-04
Q14789 (Q14789) GIANTIN (GCP372) (MACROGOLGIN) (GOLGI AUTOA	48	2e-04
Q9ZU69 (Q9ZU69) PUTATIVE VICILIN STORAGE PROTEIN (GLOBULIN-	47	3e-04

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 2 to 976 of CT614: this corresponds to nucleotides 127217 to 128191 of the genomic reference sequence.

CT615  
 Nucleotide  
 Genomic coordinates:  
 Start: 130290  
 Stop: 129405 (SEQ ID NO: 226)

Amino Acid  
 MEYIGEQLINLLDETPEEDELQLRSSFLMIGEKKYKYEVEVMSTFEAVETIRKSEFRDG  
 VFIVQLKENKHITFEGLKELRELTGDNLSLKIESLLSSIKPEKGHVILKNTSTTTDDEWL  
 ASQDKDVQEVNKLKVEKTRMLFRGFYFSPTYRYITKSLPQIPFGEKERFVVSTDFLIGLG  
 FSADDVMEKLIAIEGNMRKSGLYTWVPVAEVCHLKKYKGDIVVNPIFKSYHSHCLVIPL  
 VYLGVMFSRNVQPPSLEVETYLALAFADLYGREEMRKSCMLCEDISEVKRG  
 (SEQ ID NO: 227)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q13779 (Q13779) APOLIPOPROTEIN B-48 (FRAGMENT)	36	0.34
APB HUMAN (P04114) APOLIPOPROTEIN B-100 PRECURSOR (APO B-10	36	0.34
Q13788 (Q13788) APOLIPOPROTEIN B-100 (FRAGMENT)	36	0.34
P96470 (P96470) IGA-SPECIFIC METALLOENDOPEPTIDASE PRECURSOR	36	0.44
O28789 (O28789) SIGNAL-TRANSDUCING HISTIDINE KINASE	35	0.58
CAB55172 (CAB55172) HYPOTHETICAL 77.9 KD PROTEIN	35	0.76

#### Comments:

EST confirmation of the predicted transcript:  
 An isolated EST has equence identity to nucleotides 1 to 725 of CT615: this  
 corresponds to nucleotides 129499 to 130223 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=296  
 5'stop=318  
 3'start=360  
 3'stop=383  
 5'primer=TTAAGCCTGAGAAAGGACACGTT (residues 296 to 318 of SEQ ID NO:226)  
 Tm5=58.14  
 3'primer=TGTACGTCTTTGTCTTGAGAAGCA (residues 360 to 383 of SEQ ID NO:226)  
 Tm3=57.87  
 probel=CTACTGATGACGAGTGGC (residues 341 to 358 of SEQ ID NO:226)  
 probelstart=341  
 probelstop=358  
 direction1=Forward  
 Tm1=68.84  
 score1=1.84  
 length=88

## CT616

Nucleotide

Genomic coordinates:

Start: 147517

Stop: 144748 (SEQ ID NO: 228)

## Amino Acid

MESIKEQQQQPTVTFSEEPDQVYEFEDTTTSAKKPTPSKAKFAAGRRMVSKQRRNTIR  
 SPHTETVEEVVGEEEQQQTPEITPAEKKQOSLQELDALMGKVPALHDVSVLAKSVAE  
 FLENDEDEDEELEKNKKAQKSVLFNSVMNSGRTELSPTFCDCVSKVKSFAEGKDLVSN  
 IVKVEGEAVKKTATDTTKLANLFLGCMNLQFHEHVTIETLNKKALDKGGPLFTLKLSD  
 AVYVDEMDELEKKRQIFGSNGDKSLFKELGGNYIDSAIKSTGLVMSTPSSSSTKKAGTHFK  
 TTNQIVEESVTESMRNGCCCFKNDKWLAKRESNLKSLNNTVFGEEDDEKSAYAYSDEDE  
 DEDENEEEVDYDYNNETIESSVGNVIKNLIRKTIGLSDVEEKEEKEGEQSEEEEDSDDDD  
 DDASSVCSSSSSSVTVVAAAEDEEDDKDKDTATVVEDEDDKESVISSSEDSEE  
 DEDDDGATSCQSEVVFGDVTECEFDSDGNPLYLASDNSFRPSASVTKYPOSEEEMDVSL  
 LSKNRSTPVCLSLCRHSSGCITNSFNMSTILKSLKLFPAAGTEAAEDCVHIESTKKKDEDE  
 DEEDQGLDLQNSQYYSVLVDVNLIIIFSMGSTTYESSMVEVDYDKSFWSSFDKSVKPYCE  
 SKKSALINALCEDNVTAKVYATVHTLAIPFCESMPINHINNTTPYGSYKTRISLPGNFS  
 GQHNDINNWRSDMYTKMVENLLKREVVENKTHSRRYVRNLIVDGGVGENSGNYLKVHEN  
 NEDIFGSI EANSMSAKTAAAFKNVAKKCDLIQTNTDILTPFKQYLI DYKYNSARKNI  
 IMEPCGEDTTAHEMKRAQDAYKQALHRAKITASSISLRGIWHEMITRDMNTTYSMFMFY  
 IPDFYKYVQVSPVNVSPLYMLD  
 (SEQ ID NO: 229)

## Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q9YTL7 (Q9YTL7) ORF 48	85	1e-15
AAD56625 (AAD56625) NUCLEOLIN-RELATED PROTEIN NRP	77	6e-13
NUCL_CHICK (P15771) NUCLEOLIN (PROTEIN C23)	76	1e-12
GARP_PLAFF (P13816) GLUTAMIC ACID-RICH PROTEIN PRECURSOR	75	2e-12
IE68_PRVKA (P24827) IMMEDIATE-EARLY PROTEIN RSP40	75	2e-12
NUCL_MESAU (P08199) NUCLEOLIN (PROTEIN C23)	75	3e-12

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 476 to 1 of CT616: this  
 corresponds to nucleotides 145115 to 145590 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=1507

5'stop=1526

3'start=1554

3'stop=1575

5'primer=GAATTCGACGAGACGATGG (residues 1507 to 1526 of SEQ ID NO:228)

Tm5=59.36

3'primer=TGAAGCAGATGGTCTGAAGCTG (residues 1554 to 1575 of SEQ ID NO:228)

Tm3=59.21

probel=CTACCTGGCTTCAGACAA (residues 1536 to 1553 of SEQ ID NO:228)

probelstart=1536

probelStop=1553

direction1=Forward

Tm1=69.02

score1=1.97

length=69

CT617  
 Nucleotide  
 Genomic coordinates:  
 Start: 148612  
 Stop: 147766 (SEQ ID NO: 230)

Amino Acid  
 MSPVISQSSPSATSTAAARIISTANLRVLGVKNKEEKDEEEQQEVEPEIIEPATDFEI  
 PFSPALTICIIYINANRIHINSKGVCLNRKKIKPTSTINKNQDVPELANASSYLQTEHV  
 TDKFLSSHCSICNYVNDGEYKSALSTTRNGDQPLMRKSVRYVPLNEDNVVVQKGTYYGT  
 TFIPEKTGRRILWFSHYKKSPPITAKLCCLLETINSFNGSCSSSSSSASSSSNAPGPIEE  
 FQVSSSIFFKKEECPLQMKWVEQNELDAESPVLVLLMLAL  
 (SEQ ID NO: 231)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q9Z2U2 (Q9Z2U2) ZN-15 TRANSCRIPTION FACTOR	36	0.24
Q20150 (Q20150) F38B7.5 PROTEIN	35	0.70
O94002 (O94002) SEC12 HOMOLOGUE	34	1.2
Q20497 (Q20497) F47A4.2 PROTEIN	34	1.6
O13779 (O13779) HYPOTHETICAL 71.2 KD MEMBRANE PROTEIN C17G6	33	2.1
P93002 (P93002) REGULATORY PROTEIN NPR1	32	3.5

#### Comments:

EST confirmation of the predicted transcript:  
 An isolated EST has equence identity to nucleotides 1 to 741 of CT617: this  
 corresponds to nucleotides 147819 to 148559 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=363  
 5'stop=383  
 3'start=446  
 3'stop=467  
 5'primer=CGACAAGTTCCTTTTCATCCCA (residues 363 to 383 of SEQ ID NO:230)  
 Tm5=58.62  
 3'primer=ATCAAGGGCTGATCTCCATTTC (residues 446 to 467 of SEQ ID NO:230)  
 Tm3=58.15  
 probel=CGGCTCTAAGCACAA (residues 428 to 445 of SEQ ID NO:230)  
 probelstart=428  
 probelstop=445  
 direction1=Reverse  
 Tm1=68.92  
 score1=1.92  
 length=105

CT618  
Nucleotide  
Genomic coordinates:  
Start: 150145  
Stop: 148675 (SEQ ID NO: 232)

Amino Acid  
MESVRDVKFYTFMNVLAEKAKKIQRNLNKDKGWRTSINAEIGYGGARLMDVRFTRGRKSMDE  
LARCLYNCDGEYTTLRLVGSSAGNIIVYSLAFIMGIRGECCGFNVNNRLRMGKIIDRELF  
YKITGLNFPETVKCTCDGVRAICDLFLEVAALQEHPAWHETKEVGGKQQQHFNEFGSQYP  
GTKFNKRHLSTKIIQQMFSEEKTMEQVLAFSEGTAASGFSDLYVEAPIQYVVNMYRAIS  
NMEGRVGGAMYNLSRVLILLCSRWEKKPGYKNDFYSKCEMYIGSKKIVDDSFIFDTLITG  
DLVPLVRLAPSNEDIQRDVIRFNDSTDILMDSIDVRDVLPVLSKIIWQNV SARLKLNN  
KSLSKLAKWKWNGMVSTHDNFDSDNDYVIEHKRQLAADIMSDSLSKNHLPNFSKTITEYDE  
KENKTTPLICWNYIFELSPMGKHLFPLEEVCGFYEASLPLITPWQLKVQKKRGRQMVIIY  
GPRKRPRTO  
(SEQ ID NO: 233)

Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q64568 (Q64568) ATPASE (PLASMA MEMBRANE CALCIUM ATPASE)	37	0.28
ATC3_HUMAN (Q16720) CALCIUM-TRANSPORTING ATPASE PLASMA MEMB	33	5.4
Q12019 (Q12019) HYPOTHETICAL 559.3 KD PROTEIN	32	7.0

Comments:

TaqMan Primer/Probe Sets:

5'start=676  
5'stop=701  
3'start=823  
3'stop=849  
5'primer=GAAGCACCTATACAATACGTGGTTAA (residues 676 to 701 of SEQ ID NO:233)  
Tm5=56.76  
3'primer=GGAGCCTATGTACATCTCACACTTACT (residues 823 to 849 of SEQ ID NO:233)  
Tm3=57.26  
probel=AGAGTAGGTGCCATGTAT (residues 733 to 750 of SEQ ID NO:233)  
probelstart=733  
probelstop=750  
direction1=Forward  
Tm1=69.06  
score1=1.93  
length=174



CT619  
Nucleotide  
Genomic coordinates:  
Start: 169814  
Stop: 165116 (SEQ ID NO: 234)

## Amino Acid

MSTTQTQTIERPLPGKNNEDNSRLACLLAEGLOQQQQQDGDSEISLPLVNAGTFACYDS  
TLANLTEGRGSETENAKIRVKIHPSVFIETNKEMTIEEISTKSLNALVEKRAREARRF  
SSLTEQKFPRGGGCGYSRKNERFIEGEINNIKLNMEETASSLERLAGLLPVVINIKDWTM  
HDEKEIRLDLKGNDGMEELVNISHLNQEEWEMERLSSSIVLKDAYGVFYAHHGILDIVLT  
TSRFTGKLLQHPVIFRLMDVKVWINTPLQIAFPDTSKNPNACKILYQHPSLTRLRDLNDM  
ASNSKSVSSIIPELSKFNSTEFGMHYFTAQCFGKNTNSLKDLVTRYQLSFKNKPQPK  
LYEPTATATAASSSSSTASLTTEQKEKIAQSISSKGKSLGDVSSTLSKEYDENRKRTRK  
QKTSTDNTNIVPSGAPTSISMKNPVTCTFFGPQYTSIMDCISEKTDWIEMHLFTSLNDAEH  
NKTLVVDRKSNVSEIHDSGRFLTFGQNNTTAFIPDVIDIPTLKLILRDDSGESSAAIIASL  
IYYNNVNLEGREFSNVSDAVVGLFSGGSAITVGDIAREIASIYNIGRESNCDSILFFGEP  
ILAGRRSYGRQYRWYDPINCVVGLYRSCLETMTNRNIMRGQPVKVDETAWMYMHQQVLQVV  
LLPFFDCVLKSGVWAVKEARQLTDYIVREVLLKYTADPDQHKFLLFKKPVMDLIAKIVTH  
YAVIHSAADNGGVCIAFPDPPTFIVENDTSLRYTLDTPQSILNGDNVAENLKSATSVA  
SSPSSSSRYSETPIRVVNLVPVPTGRFLKMNKDLELFINVPLISSKEQKQQQQQTATAP  
FSSETISKSFNLVYPPKSLTRNVTYGQNIADGFLGLKNKGEIVSYFKVVKNTERRDGGIK  
DMEIGDINNHDGDTGSLSSSSSFVDGVRTSFSVDGKIEHVSALPGTTSQPTNLPVHAS  
KQVKYSVKELGMETVFFPELLSSAVLYEASKTKSTQHLSPMRIYKECVSPLSTGRIDIFP  
SKVGTVAGTGFEFIWKVLQYDTGLPTTLERLSPKIPSVPISGEDSKMEVIAESGKGVQNI  
IAIAADQLRGSNNIVGGGTRRAIQQQQQQQQEQTAQVVPVNVPARFEPTFTEIELFLQNK  
FRNVIATIIISRMMLVSNEMKIIKEVCEHVSHIMVDGLYVALDPRKAIEEILERITAEQ  
NGITIDTGNEGYGLSLRYASSGRLFINDEASEEAAAAIGGGGALGTGRRVPVELRSILDKL  
NTIGSTTQQQQQQQROQRQANNNTVPEDIKVHNEQMOKIRDSSLTSLKLLNYIRDDGRK  
DRIKTNISETLKKYSRIPSYFIASKAQKPIPWKHTKDNINLNKIPEDLNFSPAQNLFVVP  
NPRHILDMQWLNCISIIETATRDSAIVMQSFQEQADKTTTQLEELLSQWNNIVSQVTDE  
KSPAYVSSVKLEWLNNEASRIAATRENSEKSKIYMGVQGIKIVNIDELGIVAVARSIVDVD  
FYIKMPNVWASRDWKNLIYYAVNIAATPLINNISRIGIMAASQTSVLYDSSLALIAAEQAT  
RNITM

(SEQ ID NO: 235)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
ST20_YEAST (Q03497) SERINE/THREONINE-PROTEIN KINASE STE20 {	46	0.001
YN23_YEAST (P53832) HYPOTHETICAL 52.3 KD PROTEIN IN MRPL10-	41	0.047
Q24523 (Q24523) BUNCHED PROTEIN, CLASS 2 ISOFORM (SHORTSIGH	40	0.11
NIT4_NEUCR (P28349) NITROGEN ASSIMILATION TRANSCRIPTION FAC	39	0.24
O62235 (O62235) F36F2.3 PROTEIN	39	0.31
MYSG_CHICK (P10587) MYOSIN HEAVY CHAIN, GIZZARD SMOOTH MUSC	38	0.53

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 234 to 1099 of CT619: this corresponds to nucleotides 165434 to 166299 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=1920

5'stop=1941

3'start=2022

3'stop=2045

5'primer=GCCAGTGAAAGTGGATGAGACC (residues 1920 to 1941 of SEQ ID NO:234)

Tm5=59.57

3'primer=AGTTGTCTAGCCTCCTTTACAGCC (residues 2022 to 2045 of SEQ ID NO:234)

Tm3=58.24

probel=CTTGGATGTACATGCACC (residues 1943 to 1960 of SEQ ID NO:234)

probelstart=1943

probelstop=1960

direction1=Forward

Tm1=69.03

score1=1.96

length=126

CT1020  
Nucleotide  
Genomic coordinates:  
Start: 62990  
Stop: 63659 (SEQ ID NO: 236)

Amino Acid  
MEWINQRTSREDLFNTYTGNVIRSAKQALAEKHAERGERKAWTTSAASSNFNN  
VQDDYTDDITQVSIANSVLNPFLLKRYAKLIDNLAISLPPDIEDDVIIHTRDASNSTV  
RVDGANIYFAIIDGDLGVYPKQYISDKVLCGSLNREKALFYNSSKNKWTYGCNLFNDIVD  
AAIMKHPDYKEETTSTKHIRKILGIGASEKLNITHYLNFIQ  
(SEQ ID NO: 237)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
P75338 (P75338) MG307 HOMOLOG	32	2.7
Q80990 (Q80990) LATE MINOR CAPSID PROTEIN L2 (FRAGMENT)	31	7.9
Q18668 (Q18668) HYPOTHETICAL PROTEIN C47D12.2	31	7.9
Q80991 (Q80991) LATE MINOR CAPSID PROTEIN L2 (FRAGMENT)	31	7.9
Q80985 (Q80985) LATE MINOR CAPSID PROTEIN L2 (FRAGMENT)	31	7.9
Q02689 (Q02689) HYPOTHETICAL 45.6 KD PROTEIN IN COI INTRON	31	7.9

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=331  
5'stop=354  
3'start=379  
3'stop=402  
5'primer=CACACTAGAGATGCCTCCAACCTCT (residues 331 to 354 of SEQ ID NO:236)  
Tm5=57.74  
3'primer=ACCGTCAATTATGGCGAAATAGAT (residues 379 to 402 of SEQ ID NO:236)  
Tm3=58.84  
probe1=CAGTCAGAGTAGATGGAG (residues 356 to 373 of SEQ ID NO:236)  
probe1start=356  
probe1stop=373  
direction1=Forward  
Tm1=69.01  
score1=1.98  
length=72

CT1021  
Nucleotide  
Genomic coordinates:  
Start: 69264  
Stop: 76212 (SEQ ID NO: 238)

## Amino Acid

MSDTGQMEENRPATQKRRPGDEEEETGSSNVPPYANFGDDATYSMYTGEGKRGKFVLEP  
PKERSVQVRVQKPPKEKEEREQSRNVTRRPGQEFQKVLQDRSRERSEKLGQNLAEKGLQ  
ERQKKYTPKVAQMTKKIIRFREGGRKFKAPQQQTSDKGAATNVLEREEIEMAAEREQPV  
EITGDTILGGLGEEDDEDMGEDELTIQHSSMAVSQPVQQIVVSSPIPPKPTRPAPDIPIQ  
EDIVGKNISQLPPLPLDDYEDDEEHLYEEVNDFLVAPPTAAAAASTRPPRPNI PPPPPP  
VVAVADETLKNLASIAALEKEAEEQRAAAVEREREVEEQRAAAAAAAAAAAAAQREADEKR  
EREAEQRAAAAAAAAAAAAAQREADEKKEREVEEQRAAAAAAAAAAAAAQREADEKKEREVEE  
QRAAAAAEREILAQQLQEMKEQMRIKEEERRKELADKEEEKRRELAAKEEEKRQEILAKE  
EQLEKLNFLGTEITSKRALEQMLEEEKASRSRASAQAIAIQAIEYEDELPAVEPQGG  
LVPMDTDLYGKMYDLNKKLEVQNNLTSAFEDVNKTNEQNQLVAQSLEKSAKAIEKLTSQ  
KHLVPDDPAFMQRIITERDFS LKNLGNVYKRVLG VYFTLKRDLFKSKALITDKESRDLEV  
RLTDVSTDLRANDLNTILERLDVSVNIRSGGTLTKFTEADTALADQVPSRIEISNRSRS  
ALLPFSSAGLDTNFTNSSDKYNEIVNQLSSINEAMNILENIVPTLNQIKIDVTNLLTVS  
SSRQYAIERVYSVSRMDSEIRKFLAIMNSKISPYFKGDWTDERQRSIADSISSQIKSN  
DKIKESVATLHDIINTSRIRSNPLHKS SVLSSPDLNAVNDFRNFDIQQGSQFTYDVL  
SGQNIIDSLASKTTEKVTCLCLESII LDVIHKNALS LNPATY PAGETSMEESGSLA  
VDIRQEIGKNISDSSAELSR TLSEALQIFQQQQQQQQQQFQQQLLQQQQDQQNQQLLQQ  
QIEEQQRVQEQQQQQQDQQQQEQQQREQQQQQQQQREQQQQQQQQQQSDQFRQQL  
LQQQQQFQQLLQQQGRRRGGDDGEEREEREEGA EKDDCVRKVAESVATKYTADLTTLFQ  
REENNFQSKIASAKLGT LVFATPPSPIMNLT KLREEYSTFTTQCFSKLTAENNSIMRIFP  
ERIVEVCKSKN LNMGKYLYIIITTAQTEMEDRVKNILSGIFNQIEEF SNNVKQQQQQQAA  
SASSTNPPPPSTPSTPPVTSMQVCELDDQRTLEKAAIVEAITLANAVLQTTKSASAPST  
AAEREIALKLENGKTSIRMEKVDLSSGATGVSDQQKWIDESTSKQLEDFIAEENFVETA  
HNEMDIGLILDAKKNDPTRDANLRLVKPHGINVQSFYVYVLRWLGETDILDEDTVHPEY  
FRQYIDRNWKVEEHEREDTLKALGVSLSDTLAHIKDYSPSVKNDASKSVPFALNTLLYN  
IFAI DGGMISSLSRTAFIYRKFLRQSM TDKEVAQGPVRSQ LCEATIASLFTACSNLLRSS  
PLADKVEPRLQEKLA AAAAVDTSTGDMFRIRVCHLMYNFIVAYVNL CNNRINYTLNVLRA  
SGLANKKVAVAGKTTKGHTSSSHRFGSYDVTYDFS VLYKILQLQKQNISL LLEKGFNAWES  
CVAAMAAFTADPSLSISDADQSILFPLEGGEIVIEKHENDA EKNVDMVQELWKETALTIM  
AKELNSYNNWFIYSKDTDM EKLARVCRMIIGIVKAVLRLTNKAESLVD TNALS DIFKLPV  
IPIDDTKTLAINIVVFTLNNVIKPMVVSFKQMFRQKDDGVSSAYFSFQNIQQQKHQQTAS  
ILD AWACAPGKLTAAHVFI SGYENHIK LKKDDLWGASMKFPADGRGT VVEGW AQOYNN  
ESVLEDFTD FSIEVNAPASGLIIPDPL LSSMFGKNGGSSSSSSSKDNTIIGKGG LILNR  
QVVGQEQA PPINTSSDTKKIRRDANIEPIIGTPYSVIKASKGVSISVLD DFNEDSPEDFA  
LKTSIINDAIREIGQRM TYTRPIFDHQTKNIHYSSPKIILEGSDLKNGQ RSGQSWAPSS  
SSLTLASDWNLPSEL LLYRELATKQVEKEEEEKSEREEDKGQKLNEKLSFVVNKAIGTIQ  
QQHQY SERGGGMKRYQQHSADQASNGGIDDIELMNSKDATSMRKAKLALAVTNKIAAAAA  
RDGENSSAKPSNFGNRLDEA INPGALLLRGGGVRGGQTPQSSMLTMFRPGQTGGNSSWW  
TTNTPLIQRTTSVGNLVLVLPNLDSHPPTFN  
(SEQ ID NO: 239)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
076853 (076853) SRF RELATED PROTEIN	136	2e-30
AAD50121 (AAD50121) ADENYLYL CYCLASE	131	4e-29
077033 (077033) TRFA	120	8e-26
AAD46501 (AAD46501) LATENT NUCLEAR ANTIGEN	120	1e-25
088542 (088542) OPA-CONTAINING PROTEIN 1	118	4e-25
Q62006 (Q62006) OPA REPEAT (FRAGMENT)	118	5e-25

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 462 of CT1021: this corresponds to nucleotides 75184 to 75645 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=3774

5'stop=3799

3'start=3851

3'stop=3872

5'primer=AGCTGCTTCAGCTTCTTCTACTAATC (residues 3774 to 3799 of SEQ ID NO:238)

Tm5=57.20

3'primer=CGTTGATCATCCAACCTCACAAA (residues 3851 to 3872 of SEQ ID NO:238)

Tm3=57.20

probel=CTGTTACAAGCATGCAAG (residues 3833 to 3850 of SEQ ID NO:238)

probelstart=3833

probelstop=3850

direction1=Forward

Tm1=69.00

score1=1.99

length=99

CT1022  
Nucleotide  
Genomic coordinates:  
Start: 79064  
Stop: 83375 (SEQ ID NO: 240)

## Amino Acid

MDFEGTTSSTPSKMSQLYSSVKKVAEHSFANLHDKATLASKVIKDLEGERKKMSTPKSSS  
DGQKLDKAMLDDIINEYQAVKSTADNSIESTIKEIENVLESVRRTKIESEAKNSVTSSPE  
KVFSVEDLEIYSGRVCCKGLKLNANCSRIGGKYAVSMSIKKHNVSSPENNNNQVFSEEP  
DCFMLETTYPLVGFTSTEDGNTYAVFLTGVGLERSLPKYVPVFDNAGIQTLMNTGLRM  
AKLPVLCMFGRTEYDNLEDFYITSIETQSFDEEENDARMRCHTEDLERKKRMNDAPAITP  
HVAVYDYSGDGKEQLLYMITEYENTASWCNANGVVTSDSGFSNECAISDMNDLCCFADCI  
DVTVNNEEHEERSMNIVVESDRRLFDASPSPIKTEEDGENSSSSSSSPTVPPPTPYEGNA  
VVEGEEEEEEIDEDESSKYEGSEDALVMKKLAKLSTMQMRRVKNEPALKITSGGNNSSS  
SINNEEDGDDDDAVDATALCPQTEATVKNSFMAPNDERTENILYETMQISLAKICNNPSS  
MSSYRVFTNKLQECNLTMDDSIRRRPTIWTEESQQFAKGLLFDEVVTSIVAHQMAQDICK  
SEIFGGMFNANSTNIKGYEGQKSLYGNKHISSCPKTNTESNVNNALFAWVKSKLHSG  
TVIPNVFSFKMASEKPSKMKRKRRTSSASSNDEHQEPSTKMMKNDEGEKVAQESSSPSSS  
STPEQQQQAGHDKETINLIPLSFIKMPRSNVNGSASYLSEIFGQRLCGLSDASSTFKRMC  
KTFEDLENEIMRSSFTRLTRYEREVTRLYEKCRSQAVDIEENEMDVLSHQGELFAEFLED  
PIAYFEEVLENIKSWSLENVNTPKRKNKYAKVLVSVNAIRRTYEEYHAFSKFVPMFLFNL  
IKRELEGDNVYTHDVHFSSTCLWYLTVMTRNRICDVLYQINNNNNNDNEETDIVEEEEEGEG  
EEDKMEESMDVEQQQVVRKGGKKGQKFNSIGDQVIRKFVKSLCENSMMVVSIAINSLISG  
ISWMNKKIPPGFLKDSSTINTLDEVSRFVFSVKINRKINGTDDKYETVFGVSTRVDSHI  
VGPFSSIPVDFSSAGLDKASCGLYVNTIDGKGILTISPKYDLSNDEDVDSTTTDKLEKDI  
LHLSKHDTEFFNINKNKVLPFYNISPSSSLTEKKKTENRKKISSGMSNNNGMCVQTPSSS  
NSVSSVSSIVAPSSSVLALSCSLSSSTKKKSIWNENMFLTSRNMWRCGFVVPKLCSEFVN  
HRHAVKLVAETAPKTKLCRNIIIDNRKIRFENGLKKVCKSVSAFTGESTYLLNKNMTATSP  
SDLNLCIYTSSSLNDPLYTCKLTHEEYQDGNALDDYGAVFVNYTFKSIKSCSSKDEADDN  
AAAADDDGTTSTSSSTDTDAAIQDFMHVMIKKIDAMKDIRGKYKKS LAKKTKKH  
(SEQ ID NO: 241)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
UBP1_YEAST (P25037) UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 1	50	1e-04
O97236 (O97236) PFC0230C PROTEIN	48	3e-04
CAB43859 (CAB43859) HYPOTHETICAL 85.2 KD PROTEIN	46	0.001
NSR1_YEAST (P27476) NUCLEAR LOCALIZATION SEQUENCE BINDING P	44	0.006
Q21000 (Q21000) SIMILARITY TO C. ELEGANS MYOSIN HEAVY CHAIN	44	0.006
Q18918 (Q18918) CODED FOR BY C. ELEGANS CDNA CM11B12	44	0.006

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 1010 of CT1022: this corresponds to nucleotides 82212 to 83222 of the genomic reference sequence.

CT1023  
 Nucleotide  
 Genomic coordinates:  
 Start: 93228  
 Stop: 94137 (SEQ ID NO: 242)

Amino Acid  
 MVSSITHLSLLFVVAVVASVVFTEGASVRVKCAVSPCPDVIDPDHRCQGRLCRRSTRG  
 GDDDDDDDDGGTFTDVGSGILGRKKRAAPPPDEEEEDDFYRKKRAAPPPDEEEEDDFYRK  
 KRAAPPPDEEEEDDFYRKKRAAPPPDEEEEDDFYRKKRAAPPPDEEEEDDFYRKKRAAPP  
 PEDEEEEDFYRKKRAAPPPDEEEEDDFYRKKRAAPPPDEEEEDDFYRKKRAAPPPDEEEE  
 DDFYRKKRAAPPPDEEEEDDFYRKKRAAPPPDEEEEDDFYRKKRAAPPPDEEEEDDFYRK  
 KR  
 (SEQ ID NO: 243)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q18401 (Q18401) COSMID C33G8	113	1e-24
Q9YTL7 (Q9YTL7) ORF 48	98	6e-20
P73032 (P73032) HYPOTHETICAL 185.1 KD PROTEIN	95	7e-19
Q43687 (Q43687) EXTENSIN-LIKE PROTEIN (FRAGMENT)	88	7e-17
Q09085 (Q09085) EXTENSIN CLASS II PRECURSOR (CELL WALL HYDR	83	2e-15
O14686 (O14686) ALR	77	1e-13

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 214 to 1051 of CT1023: this corresponds to nucleotides 93435 to 94272 of the genomic reference sequence.

CT1024

Nucleotide

Genomic coordinates:

Start: 94623

Stop: 95742 (SEQ ID NO: 244)

Amino Acid

MDNLKGFEVALKTDLTHYKTQLDRSILVFVDVVGRLYVIVNSEQTAKKEGLATRVAKQAT  
 EIQQFKDEINNKNALNTLDDIIYIFDHGGSFKRAKHKAIIEAREYSKPLRELECMFTR  
 IADMLTLTFMTVYTNIIITEFRHSSEQATNSINVTLGRLFLCDDLCNQLPKKEEEEDLKQ  
 KFTITFHANLYMLDTRLKKDLIIFKDVIQQLHVILQKDTYAVKEGVAIRCAKQMNEISQYR  
 DNLDKNYNTFSNILEIVYIFDHGGHFEEVKHKAITLTRNYLKTLMGLKCMFKRISEMLS  
 LTFLT VYTNVIAEFINASNISDREINNYLVQLVTCNELCNQLPKPKQYRPLSLIDNIAYF  
 SLSVQKHLSGFL

(SEQ ID NO: 245)

Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
AAF04635 (AAF04635) HYPOTHETICAL 43.2 KD PROTEIN	747	0.0
O96219 (O96219) HYPOTHETICAL 139.4 KD PROTEIN	39	0.051
O64554 (O64554) YUP8H12R.45 PROTEIN	39	0.067
TPR_HUMAN (P12270) NUCLEOPROTEIN TPR	37	0.20
Q99968 (Q99968) NUCLEAR PORE COMPLEX-ASSOCIATED PROTEIN TPR	37	0.20
O97291 (O97291) PFC0960C PROTEIN	37	0.26

Comments:

Hit to public SBV sequences:

gi|6165655|gb|AF099142.1: CT nucleotides 1 to 1119 match nucleotides 3220 to 2102 of the public sequence with a 100% homology, a score of 2218 and an Eval value of 0.0

TaqMan Primer/Probe Sets:

5'start=478

5'stop=499

3'start=576

3'stop=598

5'primer=TTGTGTGACGACTTGTGCAATC (residues 478 to 499 of SEQ ID NO:244)

Tm5=58.28

3'primer=AATCTTTCTTTAGGCGTGTGTCC (residues 576 to 598 of SEQ ID NO:244)

Tm3=57.55

probel=TGCGAACCTATACATGCT (residues 558 to 575 of SEQ ID NO:244)

probelstart=558

probelstop=575

direction1=Forward

Tm1=69.00

score1=1.99

length=121



CT1025

Nucleotide

Genomic coordinates:

Start: 95824

Stop: 97369 (SEQ ID NO: 246)

Amino Acid

MDSCCLISRITPELAGKLTWIFIPENNFKIVQNSLPDDQVISQFRYFDHRHCYTFMEILM  
 ANIKIQDRKQNTTAICELTTGREGLLCRRTPVFLGSEEKREELLGNLPEGAEIFRPREV  
 MQVIGTLLDKKLEIDDDGIASVKAALCAGSSSLYLIMSHIVKMTFSAITNMKDINEEYFVD  
 FIFRHKQFLNPEFFKHLISLLKNSRKEHVAHLVRRLEHFLMLWTLSKMRFTEMEENYFPI  
 SSDSDYGICEKCAKTPKYKLRIFRERKCCDRCCRLYHQPPPEVYNWDGKITQQSNKGY  
 INAGDEIIGMLNSNDKGKTFPPPIKVVRRVVDGVYGGTILSKILKFRQANIPTCLFVT  
 CNKCNRIERLTLILGPTRNLCPPCRKKSVAVNTQQKGENKPSFVQKGTKRLRVDTGSNKN  
 TLEKFCSWERFNTVEVLLPWLGYTIESKWQNWESFLGYSSSTRYKELWAFVKNQEISSMKDS  
 YIKIEDIDQLLRSLQDQKGVFETVCKIKSRDGL  
 (SEQ ID NO: 247)

Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
AAF04634 (AAF04634) HYPOTHETICAL 45.2 KD PROTEIN	689	0.0
STA5_MOUSE (P42230) SIGNAL TRANSDUCER AND ACTIVATOR OF TRAN	32	7.5
AAD20715 (AAD20715) PUTATIVE DNAJ-LIKE PROTEIN	32	7.5
CAB60246 (CAB60246) HYPOTHETICAL 80.9 KD PROTEIN (FRAGMENT)	32	9.9

Comments:

Hit to public SBV sequence:

gi|6165655|gb|AF099142.1: CT nucleotides 1 to 1545 match nucleotides 2019 to 477  
 of the public sequence with a 99% homology, a score of 3031 and an Evalve of 0.0

TaqMan Primer/Probe Sets:

5'start=672

5'stop=693

3'start=765

3'stop=786

5'primer=CCTTTCCAAGATGAGGTTTACA (residues 672 to 693 of SEQ ID NO:246)

Tm5=57.77

3'primer=ACGGAGCTTGATTTGGGAGTT (residues 765 to 786 of SEQ ID NO:246)

Tm3=57.84

probel=TCCAGCGATAGTGATTAC(residues 721 to 738 of SEQ ID NO:246)

probelstart=721

probelstop=738

direction1=Forward

Tm1=69.03

score1=1.96

length=115

CT1026  
Nucleotide  
Genomic coordinates:  
Start: 97547  
Stop: 98789 (SEQ ID NO: 248)

Amino Acid  
MESIKLFTVAGLNMEQANQVAEEIKSEYKTEEEKRIAQEVFDKFTKKLIMQVDTSKHLLT  
RENPNRFVSRPIVHEDLWEMYKKEVACFWTLEEIDFERDPKDWEKLTQDEKDFILQILAF  
PASSDGIVIENLTTRLRQVAQIPEARSFFDFQVGMESIHGNVYGELIDRLVPDEKDKAIL  
FNAAQHFPAIKKKEQWAINWMQSNNDLAELIVAFAAVEGIFSGAFASIFWIKNRGILPG  
LTSSNEFISRDEGLHRDFACMLLKKGFVDTPSRERILEIVTEAVRIEQEFLTVSLPVKLV  
GMNCKLMSQYIEFVADKLLVEMGLEKHYNVTNPFPFMDNISLENKTNFFEKRVAEYQRAQ  
VMASINKIKKDDQTQETGSPLPILTAPPPVSSSSSEQEDVEDGVGDYISYDDF  
(SEQ ID NO: 249)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q9XYN8 (Q9XYN8) RIBONUCLEOTIDE REDUCTASE R2 SUBUNIT	388	e-107
RIR2_HUMAN (P31350) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M2	381	e-105
RIR2_MESAU (Q60561) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M2	381	e-105
RIR2_MOUSE (P11157) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M2	378	e-104
Q27124 (Q27124) RIBONUCLEOTIDE REDUCTASE SMALL SUBUNIT	377	e-104
RIR2_BRARE (P79733) RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M2	376	e-103

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has sequence identity to nucleotides 1 to 959 of CT1026: this corresponds to nucleotides 97786 to 98744 of the genomic reference sequence.

#### Hits to public SBV sequences:

gi|7672988|gb|AF144620.1: CT nucleotides 1 to 1242 match nucleotides 1 to 1242 of the public sequence with a 99% homology, a score of 2454 and an Evalule of 0.0;

gi|6165655|gb|AF099142.1: CT nucleotides 1 to 299 match nucleotides 299 to 1 of the public sequence with a 99% homology, a score of 585 and an Evalule of 1e-169

#### TaqMan Primer/Probe Sets:

5'start=642  
5'stop=663  
3'start=706  
3'stop=727  
5'primer=TGCTGCAGTTGAAGGAATCTTC (residues 642 to 663 of SEQ ID NO:248)  
Tm5=58.19  
3'primer=AGGTGAGACCAGGCAAAATACC (residues 706 to 727 of SEQ ID NO:248)  
Tm3=58.13  
probel=TTAGTGGTGCATTTCGCAT (residues 665 to 682 of SEQ ID NO:248)  
probelstart=665  
probelstop=682  
direction1=Forward  
Tm1=68.98  
score1=1.98  
length=86

CT1027  
 Nucleotide  
 Genomic coordinates:  
 Start: 98874  
 Stop: 99789 (SEQ ID NO: 250)

Amino Acid  
 MNLLPIFLTTFVAVDACSCSTICLLPDGKKQPLVFDVLEEVVYPTDVC GPKGAGELFT  
 GVDLLTLCIGGKNNNGGEWSGKGPCPRINNAVVERDYSLDEEDCKGFRKGFRIPGTDHFHT  
 VFSLCWVDRDMHAKWVRNKNINPGIVTDDLDVDSGIRTKFKYSSKIFGKGFNPRPLYSLD  
 YQERIKILKSHFNKRTGNFFARGHLAPAGDFFLASERWATFALENAPQIQNHNNNGEWD  
 IENRARTTPGAAWAETGPIFYQHKKKEYLDKKKKYIPIPHALYKIVYDKNNKELFRVQSD  
 MSWK  
 (SEQ ID NO: 251)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
CAB55635 (CAB55635) DEOXYRIBONUCLEASE I PRECURSOR (EC 3.1.2	59	6e-08
Q9Y2C4 (Q9Y2C4) ENGL-A	48	7e-05
NUCG_MOUSE (008600) ENDONUCLEASE G PRECURSOR (EC 3.1.30.-)	45	0.001
NUCG_BOVIN (P38447) ENDONUCLEASE G PRECURSOR (EC 3.1.30.-)	45	0.001
O73911 (O73911) K123 PROTEIN PRECURSOR	43	0.004
NUCG_HUMAN (Q14249) ENDONUCLEASE G PRECURSOR (EC 3.1.30.-)	42	0.005

#### Comments:

EST confirmation of the predicted transcript:  
 An isolated EST has equence identity to nucleotides 1 to 808 of CT1027: this  
 corresponds to nucleotides 98914 to 99720 of the genomic reference sequence.

CT1028  
Nucleotide  
Genomic coordinates:  
Start: 103891  
Stop: 104680 (SEQ ID NO: 252)

Amino Acid  
MSLAVTEDYGHNEKLIKRLQTSVYHTPLLGDHVMKSISDYIISRREMNNTNLLKQVEYV  
FDEETGAVIANICLLKILERCAQKGGIYDAPEDVAFFNSKMGEVTRLFTIIGGRPNMTVR  
VNEFKHGQTNNPAYGYLTDDNDTTVTTPVTPPPSPAARRSPFFTRLISESSSVVDHYVLM  
HDNPKRSSFKVYDIHAETFFPKAPSVPTFPPKTSFEISDVTLDSCMEIFSRDRDVLNVH  
DYIANDPVPFLVDVVRGSSLR  
(SEQ ID NO: 253)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
PAXI_CHICK (P49024) PAXILLIN	36	0.39
Q24459 (Q24459) POLYCOMBLIKE NUCLEAR PROTEIN	35	0.67
LA17 YEAST (Q12446) PROLINE-RICH PROTEIN LAS17	34	1.5
Q14676 (Q14676) KIAA0170 PROTEIN	33	2.6
Q94399 (Q94399) ZK265.2 PROTEIN	32	3.4
O88699 (O88699) HOMEODOMAIN PROTEIN	32	3.4

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=390  
5'stop=413  
3'start=457  
3'stop=477  
5'primer=TCCTGCCTATGGTTATCTCACAGA (residues 390 to 413 of SEQ ID NO:252)  
Tm5=58.77  
3'primer=TCTTCTTGCAGCTGGAGATGG (residues 457 to 477 of SEQ ID NO:252)  
Tm3=59.15  
probel=CCTCCTGTTACTCCTCCT (residues 439 to 456 of SEQ ID NO:252)  
probelstart=439  
probelstop=456  
direction1=Forward  
Tm1=68.94  
score1=1.94  
length=88

WO 01/38351

178/201

PCT/US00/28888

CT1029  
Nucleotide  
Genomic coordinates:  
Start: 104759  
Stop: 107330 (SEQ ID NO: 254)

Amino Acid  
MASTSSSTKKRVHEEDENLIPQPKKKKSKKVLPPFPVDKYRAVDKKVNVNLIHKILDQEKDH  
LSSTELQMITCNGAREDLLKHLDEGEFNPTIEVVSMPETIYEILSSSADDDKKFVQ  
ISLSMLIHILFFADKGTMWVSNACVQNVLGNDYKVEFENIRKKYLILEDLLNGVSNHWSE  
HGPLSHMLHSSIPIVQDMLLNRLVRYFSTYDGDQFDSFIINSVLWGIDKSVLNELTQL  
ISRGVFIVSYVPMRVTPSKDSNRPONTSPQNMSALGMKLNTFSSRISVYRNNTFKKLTE  
LVHNFYDYGSKDASSSSPPPSLSDSVNTFVRLYTNFYDIFLKVISDWKMPYGGFFKKTFDVL  
YSKGLMTLSVSEYTLKKELVFLRALKEREILYKMEKRDIIICILKKSFLGFNFRLKQL  
LPLFKHFLKIEEVKHIAFVFRDYSLMCKTQKDLQSFPPIQSASLFMEFPWLAKTWIDD  
DDDEGGKGHTLLTFAIVHRYPLISQLISHPIKSLVNTTCRDKHFTPLMHLANTSIMYQC  
NTLLCLIINGAKPEFINKFNENVLHIAIENVNYGVITELRGTLSSEQIEKMNVRMRMDN  
TTPLMIALARENIVLAQLFDGLYKPKIKVRFGSSKRLRIPEFVLLKGLKESVAYLETRNI  
SYDINI IKDAVMDNSLFEEYEIAAAGLRGNCDPEADEKTMNTWNFFTKNSTKWASSIF  
QKNRQKFVKIVDGMNRTYEDSECAICLDSLDGLPSGRTTCGHCFHNVCWLSLIRMSGPN  
NGSRARGGGIKPCSRQVTCGLGKRLGVADYDIETEEERDTKNVVPVVEEGRREWRKIGVD  
RYEFLVGGVWTNEIKL  
(SEQ ID NO: 255)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q9Y577 (Q9Y577) RING FINGER PROTEIN TERF	43	0.007
AAD27914 (AAD27914) T23O15.13 PROTEIN	40	0.058
AAD56722 (AAD56722) AUTOCRINE MOTILITY FACTOR RECEPTOR	40	0.076
Q9WV59 (Q9WV59) RING FINGER PROTEIN TERF	39	0.13
O82239 (O82239) F17A22.9 PROTEIN	39	0.13
RO52_MOUSE (Q62191) 52 KD RO PROTEIN (SJOGREN SYNDROME TYPE	38	0.22

#### Comments:

TaqMan Primer/Probe Sets:  
5'start=1406  
5'stop=1425  
3'start=1491  
3'stop=1514  
5'primer=AATTCCTTGGCTTGCAAAA (residues 1406 to 1425 of SEQ ID NO:254)  
Tm5=58.29  
3'primer=TGGCTTATTAAGGGATATCTGTGC (residues 1491 to 1514 of SEQ ID NO:254)  
Tm3=57.50  
probel=TTGGATCGACGACGATGA (residues 1428 to 1445 of SEQ ID NO:254)  
probelstart=1428  
probelstop=1445  
direction1=Forward  
Tm1=68.95  
score1=1.95  
length=109

CT620  
Nucleotide  
Genomic coordinates:  
Start: 172439  
Stop: 171509 (SEQ ID NO: 256)

Amino Acid  
MDNLITNDNIILVTFSLGLAVGCSMTIGLALAMNMLVKCIDRTTTCISCSPWEKNKNKKN  
RNGSNTSESSFISHVRFNTPDKDLDISEPLKSTTYDLANVTPQVTKLVTFSGPTYASPPT  
PRPVANTPQQQTSTNKEEESVYMPMSSSSSFSSDNSLPLTPPPSPPRSNGGDYVSYV  
NGRHLKLPSNPPSPIFNIKNEEGEDDNVEEHVYEVPEVPQQSPSIQKCIQELKEMKHKK  
NTLTRSSSNNNNNAPRITQVTFKFFPNNNNMWHNVHYGNTTIVSSTPSPTFIPSPKSI  
RKLSFKRKQ  
(SEQ ID NO: 257)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q24035 (Q24035) ENA POLYPEPTIDE	46	4e-04
Q41805 (Q41805) EXTENSIN-LIKE PROTEIN PRECURSOR	43	0.002
Q40786 (Q40786) ARABINOGALACTAN-PROTEIN PRECURSOR	43	0.003
Q9XZU9 (Q9XZU9) LARGEST SUBUNIT OF THE RNA POLYMERASE II CO	42	0.005
O36428 (O36428) GLYCOPROTEIN PRECURSOR	42	0.005
YHC4_YEAST (P38741) HYPOTHETICAL 80.1 KD PROTEIN IN SNF6-5P	42	0.005

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 665 to 1 of CT620: this  
corresponds to nucleotides 171494 to 172158 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=494  
5'stop=513  
3'start=558  
3'stop=578  
5'primer=CGCCATCTCCACCTAGAAGC (residues 494 to 513 of SEQ ID NO:256)  
Tm5=58.49  
3'primer=GAAGGTGGGTTTGAAGGAAGC (residues 558 to 578 of SEQ ID NO:256)  
Tm3=58.68  
probel=GGCGGTGATTACGTGTCA (residues 517 to 534 of SEQ ID NO:256)  
probelstart=517  
probelstop=534  
direction1=Reverse  
Tm1=69.00  
score1=1.99  
length=85

CT621  
Nucleotide  
Genomic coordinates:  
Start: 173054  
Stop: 172505 (SEQ ID NO: 258)

Amino Acid  
MGESIFDAVSLATNPNPKKSNSRNKLLRELKNMRKDFPSTFLQCRMIDFHFSGDIDKH  
YCHSVNPDPVNTIFAVFLPEEDRANNPLYDSIEGVCITVEQGLCIINKSSVHEFNIL  
VSLHKDLFGEDILDGIETASREESRSIHLYLEAGQSIRTPIPRPEGTNTVNYTIVFSNQV  
TV  
(SEQ ID NO: 259)

## Top Blast Hits

Sequences producing significant alignments:

Score	E
(bits)	Value

088632 (088632) SEMAPHORIN IV ISOFORM B  
 088633 (088633) SEMAPHORIN IV ISOFORM A  
 029244 (029244) ABC TRANSPORTER, ATP-BINDING PROTEIN  
 054948 (054948) SEMAPHORIN IV HOMOLOG (FRAGMENT)  
 Q13275 (Q13275) SEMAPHORIN IV PRECURSOR (SEMAPHORIN III/F)  
 INA9 MOUSE (P09235) INTERFERON ALPHA-9 PRECURSOR

34	0.54
34	0.54
34	0.71
33	1.6
32	2.7
32	3.6

Comments:

CT622  
Nucleotide  
Genomic coordinates:  
Start: 183817  
Stop: 180274 (SEQ ID NO: 260)

Amino Acid  
MEASNLRITEGAGVLDIDNEDDINNVDYSNLYEDEEDEEEMNEDEEEEEEDYEDEDE  
DTGVRNGRNKDPSSKKQSKFVRDVTNDMYDDDDDEEEEEEEEEDEEGEGGEYDGNLED  
EEEEGDEYEDDNEGEGEDEADPALLALAAQQEDATIIPENQWKSIVNTSPVGPNRQVL  
PMLNLFLENVAMGGSAGEEQKNKEDDQQIEPVEEEEDEEEEEQEEEEEEEEQEEEE  
EKEPIEQEKNEPEKDEDAIENESVHSHRVESSPMSEGGNDDGMDYFFSSIAGGGNDNEED  
EEDEEEEGEEEEEEPAQKSEEHVETKESVQSHTEYIEEEEEYEEYEDSRHTLEDEEIS  
TMHQFNNAAPRVRRSPPPDIQECEDAVVFPPIMKETDILPQIKEPSPKAPRMFSILGSGGE  
EQYDQLNDIAPPVPSIVTFPPDNEMGEESRDIMQDSMLMPPPPPPPPPHQPPLKP  
TNILLPPPPPPPTNQSLFSNNNNNPSFLSTVVGVNNTLGGKEAERLHKTMESIILKTR  
VKTLLLETTKNLQCELVKVVFDQDENPVKPESEKVMERLKNIIAAELTMKAFLDAAVTDI  
KSAELFRKTNEKLELFQRKQIMSNPLFSAAYASTYIMGERASKIRPSTPAPSLKKVESIS  
ELNEDETSMSSSAGGVCAEGDESIAGGGGGGGGGGGVEHSSFYSNQTQANLHMEINI  
LKEDDDNQPCQTYKLGQRLAFLNLI SFKTSSAVSWRLVNMLSDIVTKASVALFGDTNK  
AQEDFEKHQTETNDVSDLSTSSKLQMSKESANIMEEMGLSIGAEICFGAISTIEKHI  
NKLCDMDVGRITIFLNIPIVLLNWPKEFTLSKDYKVLLLDSSSCSSKMAVPPIYVLNSIQ  
FDKAVDEEDEDGNGSEAEKRSEDCNMFSEKDKKEAIRRVYDNIRYGDSNDRTSLNHFFGD  
AYSGVSNNSKNSMFDLQTQGGGRFGVAYSAGSSIEHRSPIDNALNTLVNFMKDKRHL  
LSAVVIKLLKAKLSIEVYCIKYKLNQASEKYNNKKGKHGKSTSVVPMRNLMYRPSKNQDV  
SPSTPAAATAMDVPSSVSSHVGRKRTFSFSNDINSNMSSASSVYIDQESSTPSRRRTFMD  
LLNNKSSVNSLAKQVKRMKHTKYNNSSSSEDDDDDDQYE  
(SEQ ID NO: 261)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q9YTL7 (Q9YTL7) ORF 48	154	3e-36
O40947 (O40947) ORF 73	148	1e-34
Q98148 (Q98148) ORF73 HOMOLOG	138	3e-31
AAD46501 (AAD46501) LATENT NUCLEAR ANTIGEN	136	1e-30
VG48_HSVSA (Q01033) HYPOTHETICAL GENE 48 PROTEIN	133	6e-30
Q91255 (Q91255) NF-180	130	5e-29

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=1940  
5'stop=1962  
3'start=2039  
3'stop=2058  
5'primer=CGACTCCTGCTCCTTCTCTTAAA (residues 1940 to 1962 of SEQ ID NO:260)  
Tm5=57.98  
3'primer=TCCAGCAATAGACTCGTCGC (residues 2039 to 2058 of SEQ ID NO:260)  
Tm3=57.66  
probel=GTGGAGTATGTGCTGAAG (residues 2021 to 2038 of SEQ ID NO:260)  
probelstart=2021  
probelstop=2038  
direction1=Reverse  
Tm1=69.00  
score1=1.99  
length=119



CT623  
 Nucleotide  
 Genomic coordinates:  
 Start: 196292  
 Stop: 195506 (SEQ ID NO: 262)

Amino Acid  
 MSNGATISDERLILILDKIVERRGVSNLSELLIHPITKHINELLKNTVRHGDVYMKDAE  
 LDVRSRLEDIKKDCVLKAIEKQGIDVRQIITDYLAKRKLTONLVHWYRPPISCTDIDEKI  
 QQETGQVGRCSVATYNLRIGGDDGEFTRYDFSIPLGDFKITAKLFRSINDEDVDAVILVS  
 RSDVVNDVLSFEAFNRTGERVVIFNVIVEGKSKDIDIVCKSRYKHTILNGESATYAVK  
 RIKRGDTRDDILFAITAFKEE  
 (SEQ ID NO: 263)

#### Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
PYC1_YEAST (P11154) PYRUVATE CARBOXYLASE 1 (EC 6.4.1.1) (PY	33	1.9
BAA85008 (BAA85008) ORF3P	33	2.5
BAA85073 (BAA85073) ORF4S	33	2.5
PYC2_YEAST (P32327) PYRUVATE CARBOXYLASE 2 (EC 6.4.1.1) (PY	32	5.7
O27375 (O27375) ADENINE PHOSPHORIBOSYLTRANSFERASE	32	5.7
Q88444 (Q88444) GLYCOPROTEIN PRECURSOR	31	7.5

Comments:  
 5'start=282  
 5'stop=304  
 3'start=383  
 3'stop=402  
 5'primer=GGCTAAACGAAACTAACGCAA (residues 282 to 304 of SEQ ID NO:262)  
 Tm5=58.58  
 3'primer=CGTAGCAACACTACACCGCC (residues 383 to 402 of SEQ ID NO:262)  
 Tm3=57.92  
 probel=CCAATATCTTGCACAGAT (residues 328 to 345 of SEQ ID NO:262)  
 probelstart=328  
 probelstop=345  
 direction1=Reverse  
 Tm1=68.97  
 score1=1.88  
 length=121

CT624

Nucleotide

Genomic coordinates:

Start: 195503

Stop: 194651 (SEQ ID NO: 264)

Amino Acid

MSSNGDEPAVTEAEIASVEAQLGAAHHDNSWITRKSDQLKYRLGAIAYSVAKNASIKYIE  
DQVRQEINSHLTNVMTFEHLIEDAFNPVCEAI FEKGI PVVMEKVYDVNRRIMEPREDFI  
TEILKEERWRRYIPGFYHTSFSFKYNTIAFTDSSTSFSVPINDKHMLSITPPGAAQGDLI  
DLSLSFKIDSSAKTLTLEFNKSTFAGIVNRPKSVVILSNLRNSDSSDNIGDYLKRNDPI  
YISHDTNGIINPSEDSASLTIHMPEIENASDDLYIDFNLFVF  
(SEQ ID NO: 265)

Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Y687_METJA (Q58100) HYPOTHETICAL PROTEIN MJ0687	36	0.42
Q9Y717 (Q9Y717) CAGCR3 PROTEIN	33	2.8
Q83457 (Q83457) FIBRE	32	4.7
CAB53329 (CAB53329) PUTATIVE FORMYLTRANSFERASE	32	4.7
PGK_LACDE (O32756) PHOSPHOGLYCERATE KINASE (EC 2.7.2.3)	32	6.2
Q9XEQ1 (Q9XEQ1) TNP2-LIKE PROTEIN	32	6.2

Comments:

CT625  
Nucleotide  
Genomic coordinates:  
Start: 194629  
Stop: 193327 (SEQ ID NO: 266)

Amino Acid  
MALSNNGGIYIVFAVIVLVIGASIALFFAISGVGKGLHSNAKTKKSKKYKLDSKYTDDD  
EKTDNDDNNNNGGGGGTVDVINETALQRQTRHFARTLEKAEDEFFTKLADQEFDTYKSE  
NVWLIKDKITDGKVSIEPGDINVPDVGQAIADENLFDLIGNHDEVKETMDEVVAQKSTN  
ITYEQLVIDLTNILLFGTVTVDPDENGDES LQRSTDPDAEMVMLTTTPSSQLARQQQPP  
QPTPDYLARYSKELVINNIRGGFISDRDMRTWQGRMSVHVNMQRTFNVISAATNLDLQ  
VGLEPVLQKQGRAAVGGRIEKARIEFSFVVEGNRVRVYATNKTEDCFCSLLPNCYNVKA  
SDYWISSASTAKEKTYLFIANKNDETSFFYNFEEGVVEIDLDFMTIDCAPNLPFIKNLP  
RPITDNNIMVALS  
(SEQ ID NO: 267)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
O96148 (O96148) HYPOTHETICAL 57.8 KD PROTEIN	38	0.14
TR11_ECOLI (P14565) TRAI PROTEIN (DNA HELICASE I) (EC 3.6.1	37	0.31
Q43423 (Q43423) CYSTEINE PROTEINASE (FRAGMENT)	36	0.40
O15696 (O15696) ORNITHINE DECARBOXYLASE	36	0.69
MIX2_XENLA (Q91685) HOMEBOX PROTEIN MIX.2	36	0.69
BAA84604 (BAA84604) MYOM PROTEIN	36	0.69

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 351 to 1 of CT625: this  
corresponds to nucleotides 193270 to 193620 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=651  
5'stop=669  
3'start=729  
3'stop=748  
5'primer=CCCAGACGCAGAAATGGTG (residues 651 to 669 of SEQ ID NO:266)  
Tm5=58.66  
3'primer=ACCGGGCAAGGTAATCAGGT (residues 729 to 748 of SEQ ID NO:266)  
Tm3=59.29  
probel=ACAACAACAACCTCCTCA (residues 705 to 722 of SEQ ID NO:266)  
probelstart=705  
probelstop=722  
direction1=Forward  
Tm1=69.01  
score1=1.98  
length=98

CT626

Nucleotide

Genomic coordinates:

Start: 228196

Stop: 227989 (SEQ ID NO: 268)

Amino Acid

MSDMTRNIIIVGLAVVVIALSMVAFMLSVTPLTGFLLGLGVSALGVTLFGCPTMKSPGGG

NATINPVA

(SEQ ID NO: 269)

## Top Blast Hits

Sequences producing significant alignments:	Score	E
	(bits)	Value
SUGE_PROVU (P20928) SUGE PROTEIN HOMOLOG	32	0.35
Q48546 (Q48546) INTEGRAL OUTER MEMBRANE PROTEIN	32	0.35
O86349 (O86349) HYPOTHETICAL 27.4 KD PROTEIN	31	1.0
Q9Y9X7 (Q9Y9X7) 604AA LONG HYPOTHETICAL CARBON STARVATION P	30	1.8
O69867 (O69867) PUTATIVE TRANSMEMBRANE TRANSPORT PROTEIN	29	4.0
O26370 (O26370) HYPOTHETICAL 5.9 KD PROTEIN	29	4.0

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1032 to 768 of CT626: this corresponds to nucleotides 227951 to 228215 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=60

5'stop=80

3'start=151

3'stop=173

5'primer=CATGGTCGCTTTCATGCTTTC (residues 60 to 80 of SEQ ID NO:268)

Tm5=59.30

3'primer=CCTGGAGATTTTCATAGTGGGACA (residues 151 to 173 of SEQ ID NO:268)

Tm3=58.94

probel=TGTTACTCCTGCACTTAC (residues 81 to 98 of SEQ ID NO:268)

probelstart=81

probelstop=98

direction1=Forward

Tm1=68.95

score1=1.95

length=114

CT627  
Nucleotide  
Genomic coordinates:  
Start: 234330  
Stop: 233778 (SEQ ID NO: 270)

Amino Acid  
MFQKWFESFLDSSRPRLDTTCVCSVYSYFSPCRKHIKFTSHSHHEGIKIHPPSILNHNT  
SSPTSGKMCNHHHKRLYLSTDDHTRWYDKNTSCIYLEDIGGVQFMVYEFHLTPKNNQLFS  
FPVHLQIHNRNTEKTSLLVFENEEDMRVRNIHPKSKILIPVSKDTVLVENGFRYKVKIVL  
SNK  
(SEQ ID NO: 271)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
TRBH_ECOLI (P19381) TRBH PROTEIN	31	6.4
BCL6_HUMAN (P41182) B-CELL LYMPHOMA 6 PROTEIN (BCL-6) (ZINC	31	6.4

#### Comments:

##### TaqMan Primer/Probe Sets:

5'start=187  
5'stop=207  
3'start=233  
3'stop=253  
5'primer=CCCACCAGTGGAAAGATGTGT (residues 187 to 207 of SEQ ID NO:270)  
Tm5=58.36  
3'primer=TCGTATGGTCGTCAGTGCTCA (residues 233 to 253 of SEQ ID NO:270)  
Tm3=58.89  
probel=CACCACAAGAGATTGTAC (residues 214 to 231 of SEQ ID NO:270)  
probelstart=214  
probelstop=231  
direction1=Forward  
Tm1=68.98  
score1=1.98  
length=67

CT628  
Nucleotide  
Genomic coordinates:  
Start: 240139  
Stop: 239455 (SEQ ID NO: 272)

Amino Acid  
MDSLISKLENIFSIAEQDFFNADSMFMQTMLLPTDAMFTDCESPLYKNKSGGKNIVTDVG  
ESVLSSSSDEKMSFKVLSHVLRFPVLLHCNYKQTNTPLWKELYKHGKFALIGDLVLFEN  
PFHPNIPAMPFDKSPICDITGKSIIMSEVMTKELLYKLADKDIGQFFAVLNVTNPITGDS  
FLHYFAGGNTMRDGEKDKICTSADVLRIIAEITIQKTGKMPYELMKK  
(SEQ ID NO: 273)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
SYT_HAEIN (P43014) THREONYL-TRNA SYNTHETASE (EC 6.1.1.3) (T	39	0.028
SYT_ECOLI (P00955) THREONYL-TRNA SYNTHETASE (EC 6.1.1.3) (T	38	0.049
CDR1_SCHPO (P07334) MITOSIS INDUCER PROTEIN KINASE CDR1 (EC	35	0.55
Q55626 (Q55626) METHYLENETETRAHYDROFOLATE DEHYDROGENASE	32	3.7
ERG2_SCHPO (P87113) PROBABLE C-8 STEROL ISOMERASE (DELTA-8-	32	4.8
Q85056 (Q85056) SEGMENT 2	31	8.3

#### Comments:

TaqMan Primer/Probe Sets:  
5'start=246  
5'stop=267  
3'start=288  
3'stop=305  
5'primer=GCGATTCCCTGTCCTACTTCAT (residues 246 to 267 of SEQ ID NO:272)  
Tm5=58.22  
3'primer=TCCTTCCACAGGGGCGTA (residues 288 to 305 of SEQ ID NO:272)  
Tm3=58.60  
probel=CAACTACAAGCAGACGAA (residues 270 to 287 of SEQ ID NO:272)  
probelstart=270  
probelstop=287  
direction1=Forward  
Tm1=69.05  
score1=1.94  
length=60

CT629  
Nucleotide  
Genomic coordinates:  
Start: 247143  
Stop: 244950 (SEQ ID NO: 274)

Amino Acid  
MEGGDQRTKLTTPATVMGLYQSKTPGEGEGEGGGQFKIPSAIAVKSCCSKNATRRSPPSD  
SPYSLRPMKRLKKNNGEVGGKAPPPVTLRLREDYESTPYNFNRRNKKRPITIDENQFATL  
NPTYATDIIKKQQLPSVSAASVLRKHRANADTQYRKRFSHPNCAKEFSTVNLKARDYTPLS  
VLRSHVKGPKHLKSSCDT VTETNVVKRNFSSIDKWVKLEKPPCYFAVAEADTNIAAGLES  
PFHLIRQAAGLGLISDVQDVSSNYETIKQSCIDAKEKASKFLWSNNRTKQPPSSWWPVG  
GSKNLSVLDTSPLLNWNRLCKNNGKGIKTMSIDHMAKNVFKLSPGACESILEKKTTLG  
EVTAQCKKWESYRRNIPVPAHVQPEYASQVVMIGPSELYLEVKGVYYMLETGKVIKFM  
DKEMYCEFFVFETVFSHALEGRMKGAVGVRKMCVEGFCVEMDFAGISVIDVLNGDLKCKMD  
ENVVQQPNPSTSSKPAEELMDHGSLSLRMRDTLYGVRMLQATGRLPEGLQSKCKKPITD  
SISAIAIVGKMRERMLNQLPFVLVEIVNIVTRLSQQGLVNPDIKSDNIVIDGITGQPKMI  
DFGLIVPCKKYYNFKCWGTDERFFSNHPHTAPEFINSELCSETAMTFGLAYLLIDMLSIL  
IKRTADLSANSIYTNIPFLSIVSKMYDQEKTNRPAYEIAPVIGACFPFKDNIAKLFQSP  
KHSLSYKKVK  
(SEQ ID NO: 275)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q20443 (Q20443) PRK-2 PROTEIN	46	0.001
Q17735 (Q17735) SIMILAR TO PROTEIN KINASE	44	0.004
KPC1_LYTP1 (Q25378) PROTEIN KINASE C (EC 2.7.1.-)	43	0.007
O01715 (O01715) PROTEIN KINASE C	43	0.010
O01716 (O01716) PROTEIN KINASE C (FRAGMENT)	43	0.010
Q21734 (Q21734) PUTATIVE RIBOSOMAL PROTEIN S6 KINASE II ALP	42	0.017

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 1 to 933 of CT629: this  
corresponds to nucleotides 244994 to 245926 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=933  
5'stop=954  
3'start=1020  
3'stop=1040  
5'primer=CCCTCTCTTGAACCTGGAACAGG (residues 933 to 954 of SEQ ID NO:274)  
Tm5=58.83  
3'primer=GCTCCAGGGGAAAGCTTAAAA (residues 1020 to 1040 of SEQ ID NO:274)  
Tm3=58.49  
probel=TGAGCATCGATCACATGG (residues 992 to 1009 of SEQ ID NO:274)  
probelstart=992  
probelstop=1009  
directionl=Forward  
Tm1=68.98  
score1=1.98  
length=108

CT1030  
Nucleotide  
Genomic coordinates:  
Start: 108549  
Stop: 109164 (SEQ ID NO: 276)

Amino Acid  
MSSGKVTYIEVEGGLNNKYLLDGGAAICLOSNCVARKRHAGSLHDNLFKMLGFGDPYKQ  
RRGKTNSKNLAIIEDRPQLGSVSVVOHPTPEPERFCSMTFLFAQYNMGNGRKCYFPNDKEY  
VESCKKHERVHKSTEMKRLRLYYFNKCLHAIKSPAMKKYNKIIFPARIGCAAAGGDWE  
KYHASIRDFSTIIDKEVIIVSQRM  
(SEQ ID NO: 277)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q9YTM9 (Q9YTM9) ORF 36	32	4.3
Q9XAQ7 (Q9XAQ7) NUOD, NADH DEHYDROGENASE SUBUNIT	31	9.7

#### Comments:

EST confirmation of the predicted transcript:  
An isolated EST has equence identity to nucleotides 1 to 604 of CT1030: this  
corresponds to nucleotides 108546 to 109149 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=284  
5'stop=307  
3'start=387  
3'stop=408  
5'primer=GCTCCATGACATTCTTATTTGCTC (residues 284 to 307 of SEQ ID NO:276)  
Tm5=57.98  
3'primer=TTCTGTGGAAGATTGTGGACC (residues 387 to 408 of SEQ ID NO:276)  
Tm3=57.62  
probel=TGTTGAGAGCTGCAAGAA (residues 360 to 377 of SEQ ID NO:276)  
probelstart=360  
probelStop=377  
direction1=Forward  
Tm1=69.08  
score1=1.91  
length=125



CT1031  
 Nucleotide  
 Genomic coordinates:  
 Start: 109260  
 Stop: 110088 (SEQ ID NO: 278)

Amino Acid  
 MSSNRFSQLRGNEEMVGDYSRWTTVKNNRRNRQQQYSHSFRPQQQQOHQKRTSTNSPPAPP  
 PPFPIISWGALGSYSMYRLDDQCRNCDGTGYNFHSYDRKRERVRSLNNTPSEGMWRRTS  
 RSSPFLNKKKDVDEAPPPQSNQHMYP LNKYSFREYTPSSKLVNWRDPSQEKQDKILQEEE  
 ARAPTPTPQEKEPEVETKDDVVIEETAPEPEPEPAPVDPDPDIPAITATTTTTTATRHD  
 DSSTVFLRNVILSIVFWFLGVYSALFAKCIKRSKKE  
 (SEQ ID NO: 279)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
Q01823 (Q01823) ORF-3	48	5e-05
PAR1_TRYBB (P08469) PROCYCLIC FORM SPECIFIC POLYPEPTIDE B1-	48	7e-05
Q99356 (Q99356) PROCYCLIN	48	9e-05
P73032 (P73032) HYPOTHETICAL 185.1 KD PROTEIN	48	9e-05
PARC_TRYBB (Q06084) PROCYCLIC FORM SPECIFIC POLYPEPTIDE B-A	48	9e-05
Q27045 (Q27045) SCHIZONT/SPOROZOITE SURFACE PROTEIN (FRAGME	48	9e-05

#### Comments:

EST confirmation of the predicted transcript:  
 An isolated EST has sequence identity to nucleotides 137 to 680 of CT1031: this corresponds to nucleotides 109401 to 109944 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=255  
 5'stop=277  
 3'start=347  
 3'stop=364  
 5'primer=TTGCGATGAAACTGGCTATTACA (residues 255 to 277 of SEQ ID NO:278)  
 Tm5=58.44  
 3'primer=ATCTACTTGTGCGCCGCC (residues 347 to 364 of SEQ ID NO:278)  
 Tm3=58.42  
 probel=CTCCAAGTGAAGGCATGT (residues 329 to 346 of SEQ ID NO:278)  
 probelstart=329  
 probelstop=346  
 direction1=Forward  
 Tm1=68.97  
 score1=1.97  
 length=110

CT1032  
 Nucleotide  
 Genomic coordinates:  
 Start: 119056  
 Stop: 121081 (SEQ ID NO: 280)

## Amino Acid

MAGNRTQFVSSLIAKCISDVEQGMCECCGRQAQDALMTRLANLKLGDLSKETDVNLEYLRY  
 ASTPLLGEINLDKQYAAATVDINLMAHFSYAALGIESILNSIRRVVVANHQRRNNGKKPS  
 EPISRPHPLGGVEPPLSSELANAIKDFISMGALDRLNSAIVTAALCAIASERELFLREN  
 AVNYMYDVEFAERDAATTDGTVVYLSTKMDDEDEDDIIRSEILDKVSKRPAKEGIDWRP  
 TPDNSFPYQLIWGDDSVDDTVLIDLITNAIVPNIFMAKFILFICNHLRAVIRSMREILYG  
 NISSSSDNYFEDGRKWCFWLNLYNRLEWFMLVVRVIFLHKKESFSGADNVNVKRLLVV  
 VVESFPPVLLDTEWVKTNITSWPVINNSNNNSTLPVTEDTLMRLAIRTSSGARHPIDFDEI  
 NSLTAVTNRITFQSAEFCTKILLGRALDEEEAGTKMLVKS VKETGEEKDKNNTFSSFGL  
 LLKNTKNEELEINIGDNDDETDDVACWARTSSTSFIRNRTYAFKKIWGLEDAVDVELKR  
 ESDAITSFVTDKSSPLLFYVSDWSCLLHPCCKAPAIKSVWLQILKDFSQENIKTINE  
 KVQSLSEICQKSNDRFKNKKIAAEHVRSVKLLNTISNREQEALSTEHCILWLTILWKQ  
 VVQNTLNLLNFPV  
 (SEQ ID NO: 281)

## Top Blast Hits

## Sequences producing significant alignments:

	Score (bits)	E Value
O25761 (O25761) HYPOTHETICAL 88.8 KD PROTEIN	38	0.29
O75130 (O75130) KIAA0635 PROTEIN	34	3.3
Q17425 (Q17425) HYPOTHETICAL 62.4 KD PROTEIN B0024.10 IN CH	34	3.3
O33600 (O33600) PURINE NTPASE	34	4.4
Y066_NPVAC (P41467) HYPOTHETICAL 94.0 KD PROTEIN IN POL-LEF	34	4.4
Q53626 (Q53626) ATP-BINDING PROTEIN INVOLVED IN MITHRAMYCIN	33	5.7

## Comments:

CT1033  
Nucleotide  
Genomic coordinates:  
Start: 121099  
Stop: 123634 (SEQ ID NO: 282)

## Amino Acid

MFTHLTRAFRKMNNLVNRSFIDVHRVVAELSYPEFEEDVKNPESSIIYRTPISLFQNKDIV  
TIVGDYILSPKTD SFQVLYPIKKVIEHFPVIFHCTHNNAPLWVHLLDERHHRLQLSLLTY  
EIVNAKYRGIVVIPYYRRPINYQTGKSLMSKLASVKVLDILMRCSYKFISLMCMINKK  
NNTNFLHCCASKWGEVGSKMMLHIAEMFFANPTTSQHLSDASSFPDAAAEDDKGKTPAHL  
AIQEDNADALLFLISLYGAPWFDNNSYMKSALELKSNCVKVLSFAADKYEILPNINNN  
QLEPDTMCGVCATSV EEDENEGKTTLSWYQMNCKHYIHCECLMGMC AAAGNVQCPMCRE  
DVGDEVLERCPPTIFRWLKLAEERSEHNRVLFEAKKQEFYKQMEAMKPPRVVVP RRRTFLT  
PARRGERAIRIAREIATNAIAEATAQGDVNSYFPVLIDGSGEEYEEEGEEFFNSEEEALA  
FGRPFLEDEEEARQIQMRQFAELSRRGVSVNIINNDNPHRHISTVNIVQPVYGVKSPAA  
SFIYNMLKNDVFESIRSRDTRVGGERVPVMNLSNDKRALFHAASSMLCDFATETNSQIVG  
LDFQAVYDPHHISNYIETFGSPLHAYPGAVTFLDGAQDYAESIRYDNDIVSFSEMASEL  
HITEALDVFEGSLLSPLFKKIRTGKSYSNWNHLLRRRNYARDIAEEFVRVCENSLASREH  
PPVHVHPFRDGAIPILIEYIVDFIHCITWSMQVNALHCMRKYIEHENTNVHLLNLRPTD  
ERVEVLRVSQLRWSRLFNEQYNTRMSLSTKRLSLMKIFNHD LGVSKFGVYKLLDIEMYC  
FTLI  
(SEQ ID NO: 283)

## Top Blast Hits

Sequences producing significant alignments:

	Score (bits)	E Value
Q9XUM8 (Q9XUM8) W02A11.3 PROTEIN	39	0.17
YEYC_CAEEL (Q21802) HYPOTHETICAL 64.6 KD PROTEIN R07B7.12 I	38	0.22
CAB40955 (CAB40955) HYPOTHETICAL 22.5 KD PROTEIN	37	0.50
YBR2_YEAST (P38239) HYPOTHETICAL 13.2 KD PROTEIN IN ORC2-TI	37	0.50
Q47955 (Q47955) HHDA PRECURSOR	37	0.50
O82239 (O82239) F17A22.9 PROTEIN	37	0.50

## Comments:

EST confirmation of the predicted transcript:

An isolated EST has sequence identity to nucleotides 1 to 208 of CT1033: this corresponds to nucleotides 123449 to 123656 of the genomic reference sequence.

## TaqMan Primer/Probe Sets:

5'start=1206  
5'stop=1225  
3'start=1269  
3'stop=1287  
5'primer=GGAAGCAATGAAACCTCCCA (residues 1206 to 1225 of SEQ ID NO:282)  
Tm5=58.56  
3'primer=GATGGCTCGTTCGCCTCTT (residues 1269 to 1287 of SEQ ID NO:282)  
Tm3=58.97  
probel=TTGTTCTCTCTCGCAGGA (residues 1232 to 1249 of SEQ ID NO:282)  
probelstart=1232  
probelstop=1249  
direction1=Forward  
Tm1=68.99  
score1=1.99  
length=82

CT1034  
Nucleotide  
Genomic coordinates:  
Start: 123757  
Stop: 126556 (SEQ ID NO: 284)

Amino Acid  
MAADLLELAIQETIQSELEEIADTEFLNYLPHKTGICEEAAANGRPYLPPTLEMRNEVDHF  
WSQDNRKLLKLLGHFCGNLYVEAFIAGSIDAETCVGFLRSQATGLGYPLLKKLALIAREDK  
SNTTNYNLYIDRNSMMKQVFSAEIDKRPSIQNTSHTKSSPVYLKIDRRTECLALDWLD  
ASKRTAKEIGAARKVCFLQNLIVAILIPAYTETFVLDTGNELEQQVLDDAYFNAENKDKV  
DEMCVVAILSTLHNLVVRKSLPHHLYNAPFRLPPFGQHPIINIENSSFFNEDTTPILASI  
SIPSSMVIKHHTRKNSRWRCPPNNLMTAAERSIFLRGVLTVSGDYGWFSVIVGSTIMPSVL  
FYGDRKHLINTVKSNFSAITCSYWNKYMDCRSYGFETIDTPENNCGFIRAAIDCSNTD  
FHSPVTRVNNKKTSIINAVKNPFFIRHTEPKWYNKNAMCGEVLENVGVLTLEQHVRVSDEY  
MDRFGSLLLGREKKWTCNYLDRIKSLETISNNLKGKIDTMCKILETKYNYKSSSLYYKQI  
TATSDDPKMKIIASINKRRYLCNILEFAIISSEKKDEVEEDHTKTGNGGCAFSKYKKKQ  
LEPKQHLIVKVNYIEAFSLIKMLRNDCCRNKCRFKEAIRECANELVRELYRASARSYV  
HDLVLKRTNVHLTWQRPYDENANTIMSLIPKCKLHTVLYDKDSRDVKLLNFLRTRDGNYN  
PIRHSMLLELVYGEYAKDVSTVTCFEWLKWCSSKGVIKYEDFLDRYEKTGEEDKDEREFF  
RLKKCSRDHDKDIKKIENVLNSDTLYSYSLDKNVQTHASSSTVVKNDTDGKTSVMVGWDYI  
FSGKGEKTTKKRKLETIDISSDDDDDEEEDEDEGKRMKMNNCSSSIKNKSKNKNGRMC  
CTDILNVVEPSLPNTLSFNCVKSMVDVNLNL  
(SEQ ID NO: 285)

#### Top Blast Hits

Sequences producing significant alignments:	Score (bits)	E Value
062235 (062235) F36F2.3 PROTEIN	46	8e-04
AAF04442 (AAF04442) PUTATIVE TRANSLATION INITIATION FACTOR	43	0.012
060313 (060313) KIAA0567 PROTEIN (FRAGMENT)	41	0.037
Q90491 (Q90491) DNA BINDING PROTEIN E12	41	0.048
BAA82145 (BAA82145) MYOSIN HEAVY CHAIN 2B	40	0.063
097303 (097303) PFC1060C PROTEIN	40	0.063

#### Comments:

EST confirmation of the predicted transcript:

An isolated EST has equence identity to nucleotides 1 to 647 of CT1034: this corresponds to nucleotides 125649 to 126295 of the genomic reference sequence.

#### TaqMan Primer/Probe Sets:

5'start=1391  
5'stop=1410  
3'start=1468  
3'stop=1489  
5'primer=AAAATGTTGGCGTGACCTC (residues 1391 to 1410 of SEQ ID NO:284)  
Tm5=58.49  
3'primer=ACGTCCATTTCTTTTCTCGTCC (residues 1468 to 1489 of SEQ ID NO:284)  
Tm3=58.24  
probel=AACAACACGTCCGTGTTA (residues 1412 to 1429 of SEQ ID NO:284)  
probelstart=1412  
probelstop=1429  
direction1=Reverse  
Tm1=69.00  
score1=1.99  
length=99

SBV. ORF022.txt  
Nucleotide  
Genomic Coordinates:  
Start: 265664  
Stop: 266297 (SEQ ID NO:286)

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:  
5'start=230  
5'stop=252  
3'start=321  
3'stop=343  
5'primer=CAAATCAAAGGCCATATGAAACC (residues 230 to 252 of SEQ ID NO:286)  
Tm5=58.41  
3'primer=GGTGGTCTTTATTTTCATCACCG (residues 321 to 343 of SEQ ID NO:286)  
Tm3=58.41  
primerScore=0.86  
allele1=  
probel=GAGGATAGCCAGTTCTAT (residues 265 to 282 of SEQ ID NO:286)  
probelstart=265  
probelstop=282  
direction1=Reverse  
Tm1=68.99  
score1=1.99  
length=114

C003  
Nucleotide  
Genomic Coordinates:  
Start: 126709  
Stop: 127126 (SEQ ID NO: 287)

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:  
5'start=175  
5'stop=194  
3'start=265  
3'stop=283  
5'primer=CATGGTGTGCCTAAAGCAGG (residues 175 to 194 of SEQ ID NO:287)  
Tm5=57.84  
3'primer=TGGCACAATCATGGGTTGA (residues 265 to 283 of SEQ ID NO:287)  
Tm3=57.81  
primerScore=0.91  
allele1=  
probe1=AATGGAGGCATCGAAGAA (residues 212 to 229 of SEQ ID NO:287)  
probe1start=212  
probe1stop=229  
direction1=Reverse  
Tm1=69.02  
score1=1.97  
length=109

WO 01/38351

196/201

PCT/US00/28888

C010  
Nucleotide  
Genomic Coordinates:  
Start: 184094  
Stop: 184931 (SEQ ID NO: 288)

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:  
5'start=547  
5'stop=570  
3'start=655  
3'stop=678  
5'primer=AAGATGATGAAACCTGGTTCAAAA (residues 547 to 570 of SEQ ID NO:288)  
Tm5=57.75  
3'primer=CGTCATGGAAGAAAAGGAATTGTA (residues 655 to 678 of SEQ ID NO:288)  
Tm3=58.24  
primerScore=0.67  
allele1=  
probel=CTCCTTGCAAGTGCGATA (residues 598 to 615 of SEQ ID NO:288)  
probelstart=598  
probelstop=615  
direction1=Forward  
Tm1=69.01  
score1=1.98  
length=132

C020  
Nucleotide  
Genomic Coordinates:  
Start: 60122  
Stop: 60341 (SEQ ID NO: 289)

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:  
5'start=2  
5'stop=26  
3'start=118  
3'stop=140  
5'primer=CAACCTCCACATCTCCTTCTAATTC (residues 2 to 26 of SEQ ID NO: 289)  
Tm5=58.27  
3'primer=AACGCTTTGTTTAACGTGCTTTT (residues 118 to 140 of SEQ ID NO: 289)  
Tm3=58.18  
primerScore=0.61  
allele1=  
probe1=ATTCAAGAGGTCACAAACA (residues 80 to 98 of SEQ ID NO: 289)  
probe1start=80  
probe1stop=98  
direction1=Forward  
Tm1=68.75  
score1=1.50  
length=139



C028  
Nucleotide  
Genomic Coordinates:  
Start: 277009  
Stop: 277324 (SEQ ID NO: 290)

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:

5'start=293  
5'stop=315  
3'start=384  
3'stop=406  
5'primer=GGTGCAGTTTCATCCTTACCATT (residues 293 to 315 of SEQ ID NO: 290)  
Tm5=58.07  
3'primer=TCGTCATCATAGTCGTCGTC AAC (residues 384 to 406 of SEQ ID NO: 290)  
Tm3=58.72  
primerScore=0.86  
allele1=  
probel=TTCCACAACCACCACTAC (residues 320 to 337 of SEQ ID NO: 290)  
probelstart=320  
probelstop=337  
direction1=Forward  
Tm1=68.94  
score1=1.94  
length=114

C036  
Nucleotide  
Genomic Coordinates:  
Start: 115067 (SEQ ID NO: 291)  
Stop: 115380

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:  
5'start=99  
5'stop=117  
3'start=158  
3'stop=177  
5'primer=CCACTGCTGGACGCATCTC (residues 99 to 117 of SEQ ID NO:291)  
Tm5=58.94  
3'primer=CGACGGACAGTGGAGCTCTT (residues 158 to 177 of SEQ ID NO: 291)  
Tm3=59.04  
primerScore=1.25  
allele1=  
probe1=CCATCAATGAAGAAGCGT (residues 131 to 148 of SEQ ID NO: 291)  
probe1start=131  
probe1stop=148  
direction1=Forward  
Tm1=68.97  
score1=1.97  
length=79

C037  
Nucleotide  
Genomic Coordinates:  
Start: 186828  
Stop: 187047 (SEQ ID NO: 292)

Comments: EST confirmation of transcript not predicted computationally.

TaqMan Primer/Probe Sets:  
5'start=63  
5'stop=82  
3'start=133  
3'stop=157  
5'primer=TTGACATCAGACCGACCCAG (residues 63 to 82 of SEQ ID NO: 292)  
Tm5=58.08  
3'primer=GGAATTTTACTCTTCCTCCATCTCA (residues 133 to 157 of SEQ ID NO: 292)  
Tm3=57.92  
primerScore=0.62  
allele1=  
probe1=CACCCTCTAAACTCGAGC (residues 88 to 105 of SEQ ID NO: 292)  
probe1start=88  
probe1stop=105  
direction1=Forward  
Tm1=69.49  
score1=1.50  
length=95





09/03/2000 14:45:41 04-04-03

PTO/SB/01 (10-01)

Approved for use through 10/31/2002 OMB 0651-0032  
U S Patent and Trademark Office, U S DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

**DECLARATION FOR UTILITY OR  
DESIGN  
PATENT APPLICATION  
(37 CFR 1.63)**☐Declaration  
Submitted  
with Initial  
Filing

OR

☒Declaration  
Submitted after Initial  
Filing (surcharge  
(37 CFR 1.16 (e))  
required)**Attorney Docket Number**

CL000895USNAT

**First Named Inventor**

Xun XU et al.

**COMPLETE IF KNOWN****Application Number**

09/914,464

**Filing Date**

11/08/2000

**Art Unit****Examiner Name****As the below named inventor, I hereby declare that:**

My residence, mailing address, and citizenship are as stated below next to my name

I believe I am the original and first inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled.

PRIMARY NUCLEOTIDE SEQUENCE OF THE SHRIMP WHITE SPOT BACILLIFORM  
VIRUS (WSBV), DISCOVERY SYSTEMS CONTAINING THIS SEQUENCE AND  
DETECTION KITS AND ANTIVIRAL TARGETS FOR DETECTION AND CONTROLLING  
SHRIMP VIRUS OUTBREAK AND SPREAD

(Title of the Invention)

the specification of which

☐

is attached hereto

OR

☒

was filed on (MM/DD/YYYY)

11/08/2000

as United States Application Number or PCT International

Application Number

US00/28888

and was amended on (MM/DD/YYYY)

(if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment specifically referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR 1.56, including for continuation-in-part applications, material information which became available between the filing date of the prior application and the national or PCT international filing date of the continuation-in-part application.

I hereby claim foreign priority benefits under 35 U.S.C. 119(a)-(d) or (f), or 365(b) of any foreign application(s) for patent, inventor's or plant breeder's rights certificate(s), or 365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent, inventor's or plant breeder's rights certificate(s), or any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number(s)	Country	Foreign Filing Date (MM/DD/YYYY)	Priority Not Claimed	Certified Copy Attached?	
				YES	NO
US00/28888	WIPO	11/08/2000	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
99124717.5	China	11/24/1999	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

☐

Additional foreign application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto:

[Page 1 of 2]

Burden Hour Statement This form is estimated to take 21 minutes to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number

**DECLARATION — Utility or Design Patent Application**Direct all correspondence to ☒ Customer Number or Bar Code Label 25748 OR ☐ Correspondence address below

Name

Address

City

State

ZIP

Country

Telephone

Fax

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. 1001 and that such willful false statements may jeopardize the validity of the application or any patent issued thereon

NAME OF SOLE OR FIRST INVENTOR : ☐ A petition has been filed for this unsigned inventor
 1-00  
 Given Name Xun  
 (first and middle (if any))

 Family Name XU  
 or Surname

 Inventor's  
 Signature

 Date 3/15/2002

 Residence: City Xiamen
CNX

State

 Country China

 Citizenship China

 c/o Celera Genomics  
 45 West Gude Drive  
 Mailing Address

 City Rockville

 State MD

 ZIP 20850

 Country US
NAME OF SECOND INVENTOR: ☐ A petition has been filed for this unsigned inventor
 2-00  
 Given Name Feng  
 (first and middle (if any))

 Family Name YANG  
 or Surname

 Inventor's  
 Signature

 Date 3/15/2002

 Residence: City Xiamen
CNX

State

 Country China

 Citizenship China

 c/o Celera Genomics  
 45 West Gude Drive  
 Mailing Address

 City Rockville

 State MD

 ZIP 20850

 Country US
☒ Additional inventors are being named on the 2 supplemental Additional Inventor(s) sheet(s) PTO/SB/02A attached hereto

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number

**DECLARATION****ADDITIONAL INVENTOR(S)****Supplemental Sheet**Page 3 of 4**Name of Additional Joint Inventor, if any:**☐ A petition has been filed for this unsigned inventorGiven  
Name JunFamily Name  
or Surname HEInventor's  
Signature Date 03/04/2002Xiamen

Residence: City

CN X

State

China

Country

China

Citizenship

c/o Celera Genomics  
Mailing Address45 West Gude Drive

Mailing Address

City RockvilleMD  
State20850  
ZIPUS  
Country**Name of Additional Joint Inventor, if any:**☐ A petition has been filed for this unsigned inventorGiven  
Name Lin-ZuoFamily Name  
or Surname PHAMInventor's  
Signature

Date

Delmar

Residence: City

CA  
StateUS

Country

US

Citizenship

c/o Celera Genomics  
Mailing Address45 West Gude Drive

Mailing Address

City RockvilleMD  
State20850  
ZIPUS  
Country**Name of Additional Joint Inventor, if any:**☐ A petition has been filed for this unsigned inventorGiven  
Name MeiFamily Name  
or Surname HEInventor's  
Signature

Date

Shanghai

Residence: City

State

China

Country

US

Citizenship

c/o Celera Genomics  
Mailing Address45 West Gude Drive

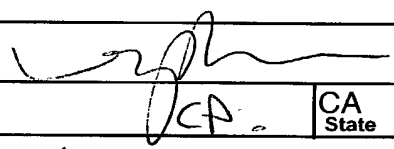
Mailing Address

City RockvilleMD  
State20850  
ZIPUS  
Country

Burden Hour Statement: This form is estimated to take 21 minutes to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

<b>DECLARATION</b>	<b>ADDITIONAL INVENTOR(S)</b> <b>Supplemental Sheet</b> Page <u>3</u> of <u>4</u>
--------------------	---

<b>Name of Additional Joint Inventor, if any:</b>		<input type="checkbox"/> A petition has been filed for this unsigned inventor	
Given Name <u>Jun</u>	Family Name or Surname <u>HE</u>		
Inventor's Signature		Date	
Xiamen Residence: City	State	China Country	China Citizenship
c/o Celera Genomics Mailing Address			
45 West Gude Drive Mailing Address			
City <u>Rockville</u>	MD State	20850 ZIP	US Country
<b>Name of Additional Joint Inventor, if any:</b>		<input type="checkbox"/> A petition has been filed for this unsigned inventor	
Given Name <u>Lin-Zuo</u>	Family Name or Surname <u>PHAM</u>		
Inventor's Signature 		Date <u>3/4/02</u>	
Delmar Residence: City	CA State	US Country	US Citizenship
c/o Celera Genomics Mailing Address			
45 West Gude Drive Mailing Address			
City <u>Rockville</u>	MD State	20850 ZIP	US Country
<b>Name of Additional Joint Inventor, if any:</b>		<input type="checkbox"/> A petition has been filed for this unsigned inventor	
Given Name <u>Mei</u>	Family Name or Surname <u>HE</u>		
Inventor's Signature		Date	
Shanghai Residence: City	State	China Country	US Citizenship
c/o Celera Genomics Mailing Address			
45 West Gude Drive Mailing Address			
City <u>Rockville</u>	MD State	20850 ZIP	US Country

Burden Hour Statement. This form is estimated to take 21 minutes to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.



**DECLARATION****ADDITIONAL INVENTOR(S)  
Supplemental Sheet**  
Page 3 of 4

<b>Name of Additional Joint Inventor, if any:</b>		<input type="checkbox"/> A petition has been filed for this unsigned inventor	
Given Name: Jun		Family Name or Surname: HE	
Inventor's Signature		Date	
Residence: City: Xiamen	State:	Country: China	Citizenship: China
c/o Celera Genomics Mailing Address			
45 West Gude Drive Mailing Address			
City: Rockville	MD State	20850 ZIP	US Country
<b>Name of Additional Joint Inventor, if any:</b>		<input type="checkbox"/> A petition has been filed for this unsigned inventor	
Given Name: Lin-Zuo		Family Name or Surname: PHAM	
Inventor's Signature		Date	
Residence: City: Delmar	CA State	US Country	US Citizenship
c/o Celera Genomics Mailing Address			
45 West Gude Drive Mailing Address			
City: Rockville	MD State	20850 ZIP	US Country
<b>Name of Additional Joint Inventor, if any:</b>		<input type="checkbox"/> A petition has been filed for this unsigned inventor	
Given Name: Mei		Family Name or Surname: HE	
Inventor's Signature: <i>Mei</i>		Date: Feb. 27, 2002	
Residence: City: Shanghai	CN State	China Country	US Citizenship
c/o Celera Genomics Mailing Address			
45 West Gude Drive Mailing Address			
City: Rockville	MD State	20850 ZIP	US Country

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number

**DECLARATION****ADDITIONAL INVENTOR(S)  
Supplemental Sheet  
Page 4 of 4**

<b>Name of Additional Joint Inventor, if any:</b>		<input type="checkbox"/> A petition has been filed for this unsigned inventor	
Yun Given Name		YE Family Name or Surname	
Inventor's Signature		Date	
Shanghai Residence: City	State	China Country	China Citizenship
c/o Celera Genomics Mailing Address			
45 West Gude Drive Mailing Address			
Rockville City	MD State	20850 ZIP	US Country
<b>Name of Additional Joint Inventor, if any:</b>		<input type="checkbox"/> A petition has been filed for this unsigned inventor	
Yan Given Name		SHEN Family Name or Surname	
Inventor's Signature		Date	
Beijing Residence: City	State	China Country	China Citizenship
c/o Celera Genomics Mailing Address			
45 West Gude Drive Mailing Address			
Rockville City	MD State	20850 ZIP	US Country
<b>Name of Additional Joint Inventor, if any:</b>		<input type="checkbox"/> A petition has been filed for this unsigned inventor	
Chinnappa Given Name		KODIRA Family Name or Surname	
Inventor's Signature		Date	
Germantown Residence: City	MD State	US Country	IN Citizenship
c/o Celera Genomics Mailing Address			
45 West Gude Drive Mailing Address			
Rockville City	MD State	20850 ZIP	US Country

Burden Hour Statement: This form is estimated to take 21 minutes to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

<b>DECLARATION</b>	<b>ADDITIONAL INVENTOR(S)</b> <b>Supplemental Sheet</b> Page <u>4</u> of <u>4</u>
--------------------	---

<b>Name of Additional Joint Inventor, if any:</b>		<input type="checkbox"/> A petition has been filed for this unsigned inventor	
Yun <small>Given Name</small>	YE <small>Family Name or Surname</small>		
Inventor's Signature		Date	
Shanghai <small>Residence: City</small>	State	China <small>Country</small>	China <small>Citizenship</small>
c/o Celera Genomics <small>Mailing Address</small>			
45 West Gude Drive <small>Mailing Address</small>			
city Rockville	MD <small>State</small>	20850 <small>ZIP</small>	US <small>Country</small>
<b>Name of Additional Joint Inventor, if any:</b>		<input type="checkbox"/> A petition has been filed for this unsigned inventor	
Yan <small>Given Name</small>	SHEN <small>Family Name or Surname</small>		
Inventor's Signature <i>Yan Shen</i>		Date <i>2002/02/28</i>	
Beijing <small>Residence: City</small>	CNX <small>State</small>	China <small>Country</small>	China <small>Citizenship</small>
c/o Celera Genomics <small>Mailing Address</small>			
45 West Gude Drive <small>Mailing Address</small>			
city Rockville	MD <small>State</small>	20850 <small>ZIP</small>	US <small>Country</small>
<b>Name of Additional Joint Inventor, if any:</b>		<input type="checkbox"/> A petition has been filed for this unsigned inventor	
Chinnappa <small>Given Name</small>	KODIRA <small>Family Name or Surname</small>		
Inventor's Signature		Date	
Germantown <small>Residence: City</small>	MD <small>State</small>	US <small>Country</small>	IN <small>Citizenship</small>
c/o Celera Genomics <small>Mailing Address</small>			
45 West Gude Drive <small>Mailing Address</small>			
city Rockville	MD <small>State</small>	20850 <small>ZIP</small>	US <small>Country</small>

## DECLARATION

ADDITIONAL INVENTOR(S)  
Supplemental Sheet

4 4

No. 8

Name of Additional Joint Inventor, if any:		<input type="checkbox"/>	
Given Name Yun		Family Name or Surname YE	
Inventor's Signature Yun Ye		Date Feb. 28, 2002	
Residence: City Shanghai	State CNX	Country China	Citizenship China
c/o Celera Genomics Mailing Address			
45 West Gude Drive Mailing Address			
City Rockville	MD State	20850 ZIP	US Country
Name of Additional Joint Inventor, if any:		<input type="checkbox"/>	
Given Name Yan		Family Name or Surname SHEN	
Inventor's Signature		Date	
Residence: City Beijing	State	Country China	Citizenship China
c/o Celera Genomics Mailing Address			
45 West Gude Drive Mailing Address			
City Rockville	MD State	20850 ZIP	US Country
Name of Additional Joint Inventor, if any:		<input type="checkbox"/>	
Given Name Chinnappa		Family Name or Surname KODIRA	
Inventor's Signature		Date	
Residence: City Germantown	MD State	US Country	IN Citizenship
c/o Celera Genomics Mailing Address			
45 West Gude Drive Mailing Address			
City Rockville	MD State	20850 ZIP	US Country

## SEQUENCE LISTING

<110> PE CORPORATION (NY);  
THE THIRD INSTITUTE OF OCEANOGRAPHY, STATE OCEANIC ADMINISTRATION,  
CHINA ;  
SINOGENOMAX CO. LTD.

<120> PRIMARY NUCLEOTIDE SEQUENCE OF THE  
SHRIMP WHITE SPOT BACILLIFORM VIRUS (WSBV), DISCOVERY  
SYSTEMS CONTAINING THIS SEQUENCE AND DETECTION KITS AND  
ANTIVIRAL TARGETS FOR DETECTION AND CONTROLLING SHRIMP VIRUS  
OUTBREAK AND SPREAD

<130> CL000895-PCT

<150> 99124717.5

<151> 1999-11-24

<160> 293

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 305107

<212> DNA

<213> SHRIMP

<400> 1

ggatccagac	atcaccccaa	caatcgacga	cccgattaca	gtagcagaag	ttggaaaccc	60
tgctcgtgagg	ataaaaaactc	cttcattgtga	tgagttgaat	aaattcatga	gattctctac	120
taattttctca	tttttaaaat	acgtcaaacg	gtttatggcg	cagagagctt	catctaatag	180
taggaatgat	atgtccgctt	tcttattaga	tttgattttt	gaaaaattga	tagagttttt	240
attatcaaca	ggatggaata	gagaagctct	gggaggaact	actggggcta	ttcttatgat	300
caggacggca	acttgtcgtc	tagcgcaatc	ttattacccc	ctccctgaaa	ccgttggtat	360
ggactttttg	caaacacctg	ggtttgacta	caataaatct	ttccctaaca	atgaaaggta	420
tatatatgaa	tttcaagcta	caatgtaaat	aattgggttaa	taaaataaag	gtatatTTTT	480
aaaaaatgtg	tttatttttc	cccaacctta	aacagatcat	tgccaggaga	aaatcgcata	540
cttaacagat	aatgcctatt	tgtatcatcg	atgtcttcgt	caaacgttgc	tccaaacaca	600
agtgtgttga	tcctattttt	cttcaaaact	gttggtttga	aatataatac	aaatttgtaa	660
gctgaaaaat	cccttcctgt	agcatttaga	cctggagaca	gtgaaacaga	gtccactgtt	720
atatccctct	ttgtattatt	gtataactct	gcatacaatt	ctaattcttt	agattcggct	780
ttgtttacat	ttagtgtgtg	gtctccgtct	ccaagtaaag	atgtgaggat	tatgtctcct	840
ttctttttta	ccagttcaga	atcggacgtt	ccttggccag	ctacatatac	ctttccatag	900
ttcctgggtt	gtaatgtgcc	gtttacaaag	ttaaagtgg	ttcctctagc	aacaccgtta	960
atggtcaagt	ttattatttc	agattcaaca	gggtaggcgt	ctttatcctt	cttgtccaat	1020
ttcttggttaa	gttctatggt	ggttacaact	atagatataa	ccacgagtat	caatgtcaaa	1080
cccgccagta	tagcggcgta	aacccccac	atgtgcattt	tgaagaaagt	tgtacaaagg	1140
tcatgagaga	aaaaaacaga	tattaaagtt	tggtatattt	tatttagtcg	tagaaatcac	1200
cactgagata	tattgtgagc	agcaactgcc	gtttcagcat	actcgaaaga	agaatagttg	1260
cgatcctttg	aagccggcat	cttgaaatga	tgtgcagtat	tatttctgct	agttttgttg	1320
aagtactgga	tagcctttac	tgagttgaaa	gtttcacggg	aacagcttga	acagttgatg	1380
gaggaagaag	ttgcctgca	ctggcaattc	acaaaggagt	gctgattatt	gttgttatga	1440
tggtacattt	tgaaggtgag	agcacgtctt	tcttcaacca	aagggtgttg	tctaattgtg	1500
ctcaggtaca	tccacgctca	cctttttatac	caacgtctta	cagttccttc	agatcttccg	1560
tcattggcttt	aatcttggtg	gttatccctc	gatctacgaa	gaatctaata	tctccctcag	1620
atctgtttcaa	aactttccac	gactccattg	ggtccagagt	ataatctgac	ctttctatgt	1680
ttttgacaac	tacaggatct	gggtgaatat	atgcagatga	tgcatcaaaa	tctacttctt	1740
cgccgctcat	gcctgtttct	agtcctgaaa	ggagtgtgtc	caaactcttc	tcttcttctg	1800
cttcttccat	ttcttctcca	ctatctaaaa	attcgtctat	atcattatca	tcaaagtcta	1860
tgtctatatc	atcatcatca	tcttccccaa	gagatgaaga	aggagatggg	gttcgtgtcg	1920

ggggaggtgg	aggagtagga	gtaggagtag	gtggaggagt	aggagtagga	gtaggagggg	1980
gtggaggtaa	caaaccttct	ccaggcacag	tccctgttcc	tccacctcct	gtattggacc	2040
cagtatttac	actcgggttg	tcaaattcaa	aaacttcctc	ctcttctttt	ccgtcttcta	2100
cagcaccttc	ctcgtctctt	ccgatcggcc	tattttcagg	ttcccttcgc	cctaatacac	2160
ctccttcggt	gtcatcttct	ttttccttct	ctttatcttt	ctttgttttt	ttattggatg	2220
ttgtagtagt	tttatttggg	cgtgacatga	aaatcacacc	aaccaccaca	cccacaaaac	2280
acacaaaaag	tatgagagat	aataaaaagg	acatcattat	tatccttatt	tttttacatt	2340
caattatggc	accagaaaaa	taaaatgatt	tgtacaaaac	cgagttaccc	ctccacacgc	2400
ctgacctctg	ccgcgaattc	ttcataccct	tccatattct	ttgctgcaaa	ggtcacgtca	2460
tgttgatgag	ctaccatttc	atatcgcata	tctcgccaaa	aagatccatc	cgcaacacgc	2520
gccgtgcat	cttcttcaat	ggcagaagtc	atgtcagctg	tcattttttt	aataagatca	2580
cgtagtattt	cctccctata	caaactctata	accttctgcc	taaagacagg	cctctctgct	2640
aaaagagaag	gattttctgt	aagctggtcc	atggtgtaat	aaagtgttca	tttatctaac	2700
ttttaaaacc	caaactcttc	taaaatatct	acagattcca	ggtaactctgc	cgccattggt	2760
gtaatcattg	tatttgcctt	gtctatcact	gcttgccttt	tatcttccac	atctcccacc	2820
tcgtcttcgt	gttttttccct	ttcttcttct	agaagccgtc	taaagacagc	ttctctaatt	2880
gaaggtatta	gccccctaac	tacctgagaa	gcttctctgt	ttacatcatt	cacaaattcg	2940
ggcgattgta	gtacgagcgt	gatcgcgctc	ttgtagatgt	cttccaaaat	gtccattttt	3000
gttgtttaga	cctagataga	gagtgggaga	cgctattaat	ttttttgttc	atgtggctca	3060
tgacctcatg	tccgatatat	gtattggttg	ggtgtaatac	aagtctctac	aaagcaatag	3120
aaggggtaaa	aacaaaacac	accatggtgc	taactctttc	ctgtacgact	aggcgctag	3180
cgtctagcaa	ggggaatttc	tctaagggaag	atgcggtgtt	ggggaaccag	ttccccattt	3240
taaagaaatc	aaacaacttg	tcaattgcaa	gacctccctc	aatagaatct	ttttctgcat	3300
cagtggaaaa	aatattcagg	gaatggaaac	aaagtggggg	agaaaaaatt	ttcgacatat	3360
ctcagaatga	agaagaatgg	atggatatca	tatccttagt	ggaaagtgt	tatgaacctg	3420
tattttctaa	atcacttaaa	cctgataaat	tggcagataa	aacatgtcta	accgcgctg	3480
cctttgcagc	actagcttct	gccgtgggat	aaaaattgac	aatcttatca	ggtagtgatg	3540
ggagtgtgct	tcaacgtaca	acaaaggtta	tgaaaaagga	ccccaaaaaa	atagcagaat	3600
ctcttttaaa	taatgaaaaa	tggacatcta	ttttattgga	cagggttaaaa	acagccaaga	3660
aacttctaag	cagacgaggt	gcactgaaaa	gcgccgaaaag	agtagaagta	cttcacggt	3720
tgaataaaact	caaggaggct	cctcttcccc	accatccgag	cctatttgat	aatttttagt	3780
gaggaaaaac	atcagcagta	tctgctggaa	cagtcacgc	atcagatatg	catttcaaat	3840
tggttgaaca	tattttttaag	gtctccttta	gaaaatgggg	tccctgtgga	gataaaactg	3900
aaagcgggga	agaagaagat	gaggaagaag	aagaagaaga	aaagaaacat	tccatatcaa	3960
gattcgtgct	tcaatttatg	aacggacaca	acgggcaaca	ttatcatagg	cccgaagtgt	4020
cttctgttta	cttttgtgat	tattatgact	atttggccta	caggaatctc	cctaattagt	4080
acaaattatc	gtcaatgcat	cctggcacat	tcaatatgga	ggattttacct	ttccgccctt	4140
tcgcagtacc	ttcaacttat	aagacagaat	tagagtacaa	aagggtttgtg	caatcaacaa	4200
atcttcccca	gctaagtttc	gactatgggg	agtttttatg	ttactgtatc	ttcggagcag	4260
attggtacaa	acacctgggg	gatgtggtag	attctctaga	aaatagttcc	atgatatcat	4320
ttgattctca	gacattgagt	ggtgtgtata	agaacactgc	taattacaaa	aggttgggga	4380
agaaaagaaa	cgaatagcc	gatttggccg	ttaggagtat	ggcagaattt	atccgcaactg	4440
aagcgcataa	ggcattgaca	gcagaagaga	tggagaaga	agaagaggaa	gaagaagcgg	4500
aagaagaagc	gatggaccag	gagcctgcag	aagtagactt	cctttcagtg	cctcatttac	4560
gccgtaaaa	tcgtcaagct	gtttctgtgt	taaaataact	tgtggagaac	gatctttcta	4620
tattggtttc	taacttcaag	aatgtgttaa	ccgatgatac	tgtatcagga	acagatacgg	4680
acaactttgg	ttctagtggg	gaatttgaag	cattatcttc	ccattttatt	ctttcaagaa	4740
tattggatga	agtgcacatt	cttaggaata	ctgatataca	aagaacccta	ttttcaacgc	4800
acgtgtctct	gtcggataaa	tctcccccta	gccgtgtccg	tggaaagcaat	gtcaacttta	4860
ataataacgc	tgggaacatt	tcttccctgc	aaacgtatgg	cggtatagaa	gagttgcctg	4920
aaaatgtact	agtcggtttg	tccggaggat	ttgaagatac	cgacatgtat	tccggagagg	4980
atgttgttgt	cgtatgggat	ggttgtgtatg	gaggaagagt	gctaagtgtc	accttcaatt	5040
gtggtgataa	ttttatccag	ctccatgaaa	aaacgcaga	aacttttaag	gatgatcagg	5100
atttagttga	acgaataaga	gatgtgcttc	agactgcaag	taagaccgga	aaccttaaca	5160
aaaaagcata	ttcaaggaag	aacatctacg	ctgttttgcg	tgaataggc	attgagcgcc	5220
ccggggacga	ttttacagaa	aaggggattg	ctctcaagga	taaaacaaat	caacccccctc	5280
cccgtgcaag	aagtgccaa	ataacggttg	aggaagtcaa	gggatttttc	agcggttttc	5340
gtgacatttt	ggagacagg	gcgctcacca	catatagtgc	agaaaccttc	agagatttag	5400
gccaaagcat	agtaaaagag	accgaaggac	tgacagctgc	aacagtggca	gaaacatcct	5460
tctctgaagg	tttagctgaa	agttaaagg	ctgatgcgaa	ttaggtcta	gaattttcag	5520
aggacgcaa	aacggttgta	ttcaaaaatg	acacctctcg	ttctttattg	gaagaaacta	5580

gggcattaag	agcaaaacaat	acttcttttt	cgctcgtttgc	aagggacatg	ggcgtccaag	5640
ttagtgcga	tttagatgct	gaatttgctg	cagagatgag	agaaacatac	cccgatgcag	5700
cccttgaaca	aaacttgaaa	gatctcgaca	aattcgaaga	gactatacca	gaaagtcaag	5760
tgaagaaact	aaagaaaata	gacagttatt	tgacagagaa	tccagaaagg	gctggcaaaag	5820
aaattaacga	cactgaactg	tcaaaggcta	cagattcagt	attggggaag	aaactaggca	5880
atgcagttac	agtgttgatg	aacaactttg	gaaaggttac	aattgtagta	ggggcctctg	5940
tgggtggccg	gttttttaggt	ccagccgctg	tcgccctggt	gcatgcgtcc	agaggggcac	6000
atctcaacgt	cgtggaccac	accagcccta	aagggtgtcat	cagttataaa	attgtggact	6060
tttcttgtgc	agatagaaac	accggtggg	ctaagccaac	caagcaccgc	ttcaggggaag	6120
aaatagacca	tgttatcgca	ctagatgcat	cattcttaac	tgaaaatgga	gcatatgtat	6180
tccctgaaga	cggaggaccc	aaatcgaaat	ataaggccta	cgcaccaatc	tgtgggaacaa	6240
aagatgctgc	tcaaggagaa	tgtggatcct	gggcaacatt	cgacgaccgc	cattctgtat	6300
tgccttggtg	ggcaagcatg	aaagatttgc	ctaaaggaca	atccctctcc	tgcgataaaag	6360
ggatgtccac	tttaaaggca	gtttcttccg	ttcttttgtc	cataggaaag	gatgttgacg	6420
aggctatatt	tgaggttgca	gaggacgccc	tgggtggggt	ggcgagcaag	gcaatttcag	6480
ctgtaataaa	taaccccttg	ttcatatttg	gagtgcctct	tggatttggg	atagctgcta	6540
cacgcctcaa	tccatccaac	tggaaaactg	gcctcattgt	attttcaata	ctactagtgg	6600
tcatactgat	agttcgcttc	tttgcagggt	cgggcccgct	aaccctgaat	tggtttgggtg	6660
caaagaattc	agctaaaagg	aaacagactg	aacaattcga	agacggggga	ggaaatcggt	6720
caaaaatagt	attggcagaa	aaggacaacg	ccaatagtaa	acttcaatcg	aggaggaatg	6780
aaactgggcc	catgagatta	gaggagcttc	ctgggcctga	agatttgccg	ccagttttct	6840
tccctgccac	aacaaattat	tccaaatctg	ccaagattct	gggctacaaa	tctaaaccct	6900
tcaacgactt	ttatacaaaa	ataataaaca	cagacatcat	aaaaatggat	aggtaaaata	6960
acacattaaa	tgtatatatt	atataatttt	attttagtaa	caataaagat	tatacatgta	7020
taaaaattgc	attgttttat	tcctatacac	caacaataga	tatagactta	tcactaagat	7080
tataagtttc	tgtatgcgtt	tttccatcac	caattctgat	tttattccta	ttcttctcca	7140
aaacagattg	cattccgtgt	ccatcaacat	ccgtgttctt	cttgaagatg	tgggttatat	7200
atggctccga	catatcttgt	actgtcgatt	ggatagtttc	agctagggaa	tgtttttctg	7260
gtccaaattc	tcctctccca	taacggatgc	gacgataagc	ttcagcagaa	gccagcaggc	7320
aaaagtgggc	ccagccagcc	tcgtcagtac	acatggggtg	cttcagagtt	tttgttctat	7380
tgtgcacatc	aacatatgcc	ataaaatcgg	aatgggtcaa	gtgcatgtcc	tttaatttca	7440
ccgtttttga	caagtaagtg	actggcaact	gtttggcagt	aaattgttca	ttgatattag	7500
aagtggagta	tttttcttgt	ttagaagcat	gtcgtatata	gagagcgaca	gctaaggcag	7560
ttattgctag	aataaagaaa	acaccaacaa	ccataggcca	attcattcgt	ttagtagtta	7620
ttacaatttt	gttccacatg	gtcataattt	ctcagaaaag	ttcctcaaca	ggggaattat	7680
agtacgtgat	agaattgcag	tgataggcac	aggcggatat	tccgacggtc	ctcttccgag	7740
aatattttgt	ctagattgaa	gcattgcagc	aatctcctgt	ttagcatgga	caacctgtat	7800
atcatcacga	ggcaaaagat	tgtcctttga	tccttttgat	aaagcttcga	tgtttataaa	7860
ggaacctaaa	agtttccaca	ttatagcaat	gtgataaaaag	ctgtccaacg	aagcatatac	7920
agcagctggg	aaagttgtcc	atattaattg	cccactttgt	aacgcagagg	ctaaattttc	7980
aaccccgatc	ttgttgcttt	ggcagatggt	cgccatccaa	tctagttttg	ttatcgtatc	8040
tagaactggg	gtattcttag	ccagtttgac	tgcctccctt	ttcggttctt	tggggagata	8100
aactggctcc	ttttcagtag	agttggtttt	tatgtttttc	tctgatgcag	attctccaaa	8160
atataacagt	tggcgaataa	tggagtcaac	ttcaatagta	tcactcttca	ctaattctat	8220
gtccctttcg	attagttggc	acacgtcttt	atctagcctc	atcctgttag	agtttccctgt	8280
ttcggctgct	ggtgctgtcg	cagacgcagt	tgtctattagt	ccatcagatg	cgtccagtat	8340
ggaatttggg	ggaggagctg	caacagataa	catggcagta	atatcatgga	tgattttgga	8400
gtagtccagt	tctaatacat	caaaatccctg	ttcttccctt	tcctccattt	tgttgggtgc	8460
catgtttgag	gagggatggt	tgttgatact	acccgagctc	atgccgtaca	ccgaatgacg	8520
agaggaggtg	ggagcaggcc	ctttttatag	ccctaaaggg	ccattatcga	cagtactatt	8580
atagaagtta	tttttatccc	tgatattata	caagcagaga	aaaatttcag	tctacttgca	8640
ggagacatct	ttttaataaa	tttagggagt	cttctccaag	ttttcttctt	attgggtttc	8700
tgtttcttgt	ctgttggttt	ttgttggttt	ttgttggttt	gcttatgtat	ttggggttgc	8760
cgattacggt	gcctcttctt	ccttttatct	ctgtcagcaa	gtaaaatata	cttttcgtac	8820
ttgataaatc	tatcaacgta	cttttttatt	agcctgtttc	tggctgtgga	attccgtgaa	8880
gaagatcgta	gaaagtttgc	atcagcagat	gtggattctg	tatctgaatc	ttcttccctcc	8940
tctctcccaa	gaagagaatc	gacgttagtg	tctgattcag	agtcatatcc	ttcttcttct	9000
tcttcttctc	catcgcttga	aggatggcc	gaagacgtgg	gagaagaaga	ggaggatgag	9060
gaagaagatg	tagaagtaga	agccatttct	ttgtttgttg	ttgttacctg	cccatgttcc	9120
tcctcgacat	catcttcttc	ttcaatctta	ttcagtgcac	ggttcaattg	ctgctgataa	9180
tgccttgtgt	ttgttacagt	gctagaccat	aaagccctcc	tgggaagcag	cgatgctcca	9240





tgttgcgcaa	aactctagta	tcttcttcoct	cccaagtctt	tgccaggtcc	tttctaaaag	12960
agtactaat	atthtgagtc	ttgtggaact	cgtcacataa	taatgctgtg	acaaagttgt	13020
caatgggcat	aacctttgtt	gaatcgaggt	ctgttctctc	tgttataaat	ttcagtacaa	13080
attccagatt	tgtgcgtaaa	tcagagtacg	tcataaaatt	aatgggtggc	gtttctgtta	13140
agaaattccc	tattgactct	gaagaacgct	tactatgaaa	ttccttaaca	tccttgagag	13200
tgaaaaatct	aatatccacc	gagtcctttg	cagacgatac	agaattttcc	acatctactg	13260
tcatggaccc	acttttcccg	tcgctcttga	aagacgttgg	tttatttagc	caaggcactt	13320
gtagttttct	ccatgaaaaa	ttaggatcac	gaagacctct	agaattttgt	accacttctc	13380
gcatggccct	ttgacagctg	ttgaaaaag	gcttagagca	cgtaaccatt	ataaaaaagg	13440
ggtgcctttt	ccacacgcca	ctgttttagt	ccgtaattac	cttgcgcgaa	tcttgacca	13500
aaattgatgg	ccctacgtgt	ttttgtacat	ctaagagtgc	cttttcttgt	agtatgcac	13560
tcatgacgaa	ggctgccaat	tctcttgttt	ttcctacacc	agtaccatct	ccgatgataa	13620
atcctccact	cgttttcttg	tccactttaa	aattcctttc	gaccttatct	ttcacaaggt	13680
ctataatgga	catgtttctc	ttccaccctc	ttcctttgtg	gaattgaatc	agggttagac	13740
gcactgtctg	tgctgcatac	gaattaaatg	ccatcggccg	aggaagaaat	ttgggttagg	13800
catcggttac	atthtttgag	aagcccaccg	accggaatgt	ttcactctca	taaagagcga	13860
atgaagggtc	tgggtgggag	agggtgccaa	tctctttagc	atcttcttcc	tccagaaata	13920
aaagtttgtg	agccatggta	attcaaccag	aaaatacaac	aataatgttt	gtaacaagaa	13980
gttgccctta	gaagactctt	gtgtactttt	tttacgatac	ttgaaaacag	tacgctgtca	14040
cggatcaaaa	atthttgtata	ctgttttagt	gctatttttg	gcaaaataat	agttttttat	14100
ttcccatgct	aaagagaacc	aggagttagc	tacagccata	ctagtctcta	caagtaatcg	14160
tctataagcc	tcagaagtat	ccacaccgtc	cctgaaccac	gaagtttctt	gcacgtccgc	14220
ctcttatgct	cccactcga	cgagtgaact	cttgaataga	gagagggtg	ttattgtagc	14280
tccttctggg	gttgaggaaa	tgtttgacac	atcaaacgct	agacgagggt	ctgactgtac	14340
taccgggcca	ctcttggggt	tatgatccaa	acgaaacgat	tgcaaacccc	acgcatactc	14400
gtcaatatcc	actctgtact	cggatgagga	agatgatgtt	ccgaacaatt	tttcaactcc	14460
gggagttaca	ttggagcaat	caaaatagct	ggggattgtt	tcgctgtcaa	atatactatt	14520
gtggtctact	gaaaaaatat	cgtcggtagc	tccttctctg	ggaagaacag	agtaggatgt	14580
ggggaaagcg	tactgcaag	gattggacac	tttcaatgta	ctggaagttt	ttgatgcagg	14640
tttctctttt	aggtaatcta	aatcaaacat	atctcgcgct	agcacagtag	cagcataagg	14700
tccgtatgaa	cgcattaata	acccctcttt	caatacaaaa	gtatcctcag	atthttttat	14760
gatatgtctt	agaaattctt	tctctttcat	taaactatcc	caagaacttg	atthtgagata	14820
tatttcaact	gtctcctatc	ctctcgcgtg	tttctttatt	acgtattctg	tctgagacaa	14880
tctcctctca	tcgttaaaat	tacgttcaaa	aaagctccct	atttcttcca	gcctttggtc	14940
acctgtagaa	ttgtactctg	gagctggagt	acctccaggg	ggagttctga	tataagagga	15000
gggtttatth	tgactcaaaa	gtgtagagaa	taaaagttca	gatccgcgct	gaacactacc	15060
cgaaaccatg	cgactatttc	gcaaacattc	tgtggccaac	atactatgtt	catagggttt	15120
ttgagggtat	tttccggcca	aaacattagc	tgctggctca	atgtaaaagg	gtgcagtaca	15180
gaatggccta	tgtctccccg	aggaaggttt	gttcaggagc	ggagctaaac	aacttctatg	15240
taacaatgaa	gacgggcgga	tgattacata	gggaattatt	tctagaggag	gtaaaacagt	15300
gctagggtatt	ggtagcaaaa	gcacattctt	agaagtattt	gttgactcca	ggtttaagg	15360
cgtgatgatg	ggcgtttggg	ccagtgcatt	tttcttggca	gtagtttctg	cgatacttga	15420
tttaggtctt	tgctggtgaa	tctggtcccc	tcttttagat	tggtcggcca	agaaaattcc	15480
catggcagga	acagtcccc	ctaaccggtg	cctgtcatct	cctattattg	catgcctggt	15540
tgagtcttgg	ggattattaa	cgcccggtac	acacttcat	gtgctcacat	caaacgtaac	15600
aaacgaagg	tctatgtttg	gcatcataac	acagggtttt	ttcatacatt	catacatacc	15660
tctgaatttc	atatactttg	ttacatcttg	agaaagaagg	tcgttattta	gtacatcttt	15720
cttattatta	cagaaacagg	tgatgggttt	tttgtataaa	aattcaggat	cgttgcattt	15780
aatthttga	tgcgccaaag	aagaagattg	acgcaatttg	gagatgtaga	aagaggttgc	15840
gaattctgta	tgactccgta	tttgttgttt	ctttaataca	tcataataga	ttgggttcga	15900
agatgaagga	gggtgtgttc	ctttgtgggt	tgaaggattg	aaagcgcaag	cagtcacgca	15960
tgttctctcc	gaacctccga	aatatgtttg	gtatttaggt	ctgcattgcc	agccccatc	16020
atctgtcaga	acgactatag	atgtgtgcac	agaacactct	ttaactacac	gttctccagt	16080
ggttggggat	gtttctatgt	ggtcacccgg	cacgttctct	gaccaacaat	aaggtttagg	16140
atcttctcca	gattttatgt	tgataaaatg	tcctccagca	acaaatgtgt	tcatacacg	16200
aactgtcttc	gaggtataac	acacagatgt	gcaagtttgt	ccctctttca	aaggaaactg	16260
gttgaacag	tacccgccag	tattgtaggc	caaaatatcc	acatcttcat	catgtctctc	

tgaggatcta	tttgagccta	aatcatgact	ttcgaaccga	gaattccctt	ccttcgcttc	16620
tgacggaacg	ttccatttat	tccctttaac	agtcaaattc	atcatgtcca	ggagagcggg	16680
gtgaggtggg	taattgcttg	catcttccac	cccgtcatgc	ctacctctgg	cttctaggtta	16740
ttcttctagt	gtcttttcag	ggttagctcc	ttgaactgct	ttcaatttac	ttttcatccc	16800
taattcgtgt	atagcttgaa	aattaacaga	cattgggttg	gctacattat	aggccattgc	16860
agcgcagact	agcgcaaaaa	tacagcctaa	cacgaacacg	gtcacgaata	acgttatattg	16920
cgcggtgaag	cgcatacttc	ggttggtagg	gtaaagtaac	ggatctcttc	gtcgttcctc	16980
catcgttcct	gatattacaa	tggatatctt	cctagataaa	tctcatacag	gtgccaccct	17040
tttagggtct	tataagacgc	agactgtgaa	aaataattct	acgttaaaaa	agcttaaaat	17100
gggcgcgcct	accaacgctg	attttacacg	cacggatatca	ggagtagctt	cttctctcta	17160
tcttggttat	cctggagctc	cttccgatag	agaaaagtta	gttttgccct	cttcttatttc	17220
tgactctttt	gtgtacaact	acaaggatgc	agtgggtgacc	gctgaggctc	ccaagtgggtg	17280
tccctttaac	gagccagctc	ttcatgagca	catcatgaac	agacttgaaa	aagctggtct	17340
aattaacaga	tctcggtttg	tgtgtaaccc	tgttaaatcg	gctggagaag	tatgccgatt	17400
tcgctattct	ggaggaagta	ctccccagaa	cttaattttc	ccgattggag	catcagagat	17460
aaattggttg	gtcattactta	gaaatgctgc	tcgttttggg	acagtggcag	catcgcccaa	17520
agacgccatc	gaacgcattc	ctgatctaa	agaaggtggg	acaagtaaac	atgtggcaaa	17580
gaatgcaatg	aggagacttc	gtgtatggcg	agcctttaac	tggatagcgg	aagcctccag	17640
atcagcgggc	atgattcgct	acgaaccctt	ttcgggtggc	tgtgctctat	acgatcacga	17700
ggtaagaaga	aggcaactaa	aaggaaagta	tgagcgaaat	gtactatttc	ttggtgacca	17760
tttctgtaaa	gaatcaactc	ttctggctga	tatttcaaga	gggggaagaa	gttctgattt	17820
ctggacgctc	gttgaagctg	tcateccggt	caagaataga	catgctcgaa	caatcagtaa	17880
tgaaactaat	gccatccctg	aagactcttc	tataaacttg	gagtgggagg	atgttctagt	17940
aaaaaaccaa	cgcgacacaa	atcaaggaga	tgattctacg	ttagaaaaga	ccttagaagc	18000
tgccatcaaa	gaacatgaaa	gcataggaga	aaaaaggaaa	aaacatatcc	tagagtttat	18060
taaaacatgt	ctaacagagg	agcaacgaga	aatgattttt	aaggagggtg	gaggaaaagg	18120
gaatttgtcc	ccggcccatc	taacaaatct	ggcgcgatgc	atactggcta	ataatgccaa	18180
agcaggcata	tgggttatta	tgcagagtct	tttaaagcaa	atcaatttct	ctatactcca	18240
tttaatatggg	tacgaagctc	aacgactttt	aatgttcaag	ttgtatatgc	ctgcattact	18300
tgctcttttt	atatctcaga	ggggaattgg	tgacgtgttt	ctgaatgggtg	tgtttaacct	18360
agaagtgatg	aaagaaagag	ctgcaaatgc	aaaaataaga	gacatgggtt	ctcgagacgc	18420
ctacaagaac	gaatacaacg	agtctaataa	tttgggcaat	tattcccaat	tcgatattgc	18480
tacaatttat	ggtcaaatgt	ctgattgtaa	tgcaattcct	ctatctatta	atattggagt	18540
acctctccat	aaatcgagga	tgaacatgca	agacattgaa	aagaccatac	aacacatagt	18600
ggatatgggt	ttggcaaaaga	tacgtgtagc	ggaagaaaat	agagtagcta	gacgtctttt	18660
acataagaaga	aaaatgcaag	aaaaggcagc	cagggagcga	gcagctgcga	gacagagaag	18720
actaagagggt	gaagaagaag	aagaagaaga	ggaggaggag	gaagaggagg	aagaaatgga	18780
agaagaggaa	gaagaagctg	gaacatcagg	tggttaatgg	tctgggttatg	atcaagaaga	18840
agaagatgat	ggtgaggagg	aggaggaggga	agaggaagaa	gaggatgaag	aagatagtga	18900
gggtgaagat	atgaatgggg	aaaataactc	aagaaaacgt	aaaaacactg	gtaatacttc	18960
ctccactcaa	caacctcccc	aaaagcgtca	gcgaggtaaa	aatgcgcccc	tttcaaccaa	19020
gggaaaaagg	ttaaaggaaa	gagataatat	tggaggattt	ttattagcca	caattcaaaa	19080
cgatgaccgc	caagtgaatg	tggagagcat	acaaaacttt	ttgacggcca	ggcaaagaaa	19140
atatgttaaa	ggtggaaaat	gtggagataa	tggactacca	gaattgttaa	ttgaaaagggt	19200
tacaaatcta	ctagattcag	tgttccattt	cagaaagggg	tctattctta	acagcataca	19260
cgcaaatagg	cgatcagaaa	ctggcgtata	caccacaaa	gcaaactgta	tttgtgatta	19320
ctatgaaaaga	aacgtatcaa	aagacagttag	taataatata	ccccattcat	ctgaatgtat	19380
tgagcgggca	agagaaaagg	atgcttccctg	tgcgaatct	aacaaacgcc	cttgtcctgt	19440
agactctaac	aacctgaag	atgtggaaca	acgtatgcgg	gaattgatca	tggaccctcc	19500
ttcattatca	ggggtagaag	attctctcgc	tatcgaaaga	gtacttcaaa	atgagatatt	19560
attcacaagt	cttgtcacca	atccaatttt	caatgccgtg	ttaggtgctg	aaaaaggaga	19620
tcttggaaga	tttatcgtat	taaataacat	tgtaaaattt	atgaatatga	ctattgcttg	19680
tctagtggac	ggagatatgc	ctatgcttct	agactcaaga	ggcaagacaa	aaaaccttct	19740
agaaaagggt	acagtgaaaa	acacgagaaa	atttttcaaa	ccaaatatga	ctgcagcaga	19800
attgaacgtg	gccactgctc	aatctgcagg	acaccaatac	atgaacgctg	gtcattgtcc	19860
agaacctgga	atcaagcaaa	tgttactccc	tgattgtata	atgaaactga	aatccatcgc	19920
catgaaaaag	ggtcgaggag	ggagatcggc	ccttcacaga	cagaaatgtg	accatgcctt	19980
ttgcaagatg	ctgaagtgtt	tattcttcaa	tattgacctt	tcaaatgctg	cagatacatt	20040
tattgacctt	gcgtcacgtg	ccaccttatt	taggcttgat	gatctttgta	gggataggaa	20100
aaaatacaag	aacattgact	gggtaaagga	tttactagac	cctgtaatga	agggaaacaa	20160
taaatgggtg	ggaactggag	aatacactaa	cattggacgt	gactcaaacg	tggccgcccc	20220

tggtgatttt	tacacaattt	tgaatacac	aatgattgat	gatgggtgtaa	taagtgttcc	20280
ttcacggaaa	ccaaatgacg	tgtattacag	tactattgaa	agagcggatg	acttactcac	20340
cgagagcagg	gacgcttcat	gtgaaagcta	ccgcccagacc	ttgtttgacg	ccagagcggg	20400
gttggaagtc	aacggagacg	gacgtgtccc	ttacccttcc	agtgaaccg	tagaagactt	20460
gggaggagaa	gaagaagaag	aggaaataac	aggcggttatt	gacgacagta	ctgaaataga	20520
agacgttcaa	gtacaggaca	gtaatctatt	cgatgtggaa	ttattcgata	tccttgaaat	20580
agaacaacat	caacagggtg	gagaagaaga	ggaacttcct	tcagcaatat	ctgaagtgtt	20640
tgcttcatta	ccagctgata	acgattccctc	ctcaccgcg	catattccct	catttgggaa	20700
ttctgaggag	ggggaaaaaa	gtccagaacc	gtacaatat	tttgattctg	ccctcgacca	20760
attattggat	ttaatagata	gcgatggaag	aaacaataac	cctaaaagag	tcgactggaa	20820
cagtgtcacc	attcaagaat	gaaaatatta	ccagaatagc	gtctaattac	gtccgagcct	20880
ttactgacac	gtggtctcat	ttggtcaaca	tttctggagc	tcctctaact	gctgagaaga	20940
atccaagcgc	tatccagct	aacgaactga	atagatactg	gaccaaact	aacgtgttat	21000
gcaacccact	ctttaaattg	gaggaccaca	taacgaggga	tgaagatact	ggtacaataa	21060
cactaaaaat	caaaatgtat	atagatgata	aaaatggact	atatcagctc	gccgttttaa	21120
tgcttgctct	cgattcgctt	gtttcgcttg	catcttttcc	ccatggagct	gatttggttt	21180
ccaataaaag	tgaaaacaaa	ttctgcgtaa	aaattcccca	cgacactcgc	gctgaatcct	21240
tactgaataa	tggtggattc	cccgtctgat	taagtgggcc	ttttaagaga	tgagcatta	21300
actacaaggc	tgcaaaactg	agcggtaaaa	gcggtataga	tgccctttca	ggttccatgt	21360
tgacagtact	aaaaaataac	accaataaaa	gagcaactga	tattttacat	ttggtgaata	21420
atgtttctgc	ctctgcacaa	caacttgacg	attctgaaat	gtctcgtact	tttaaccacc	21480
aaaagaaagt	tggagtttgt	tatgatataa	atgtgtccag	ttcaaggcaa	gtgaaccagc	21540
gtaatttact	ccaccatcag	aatataatag	gacaacatct	gatcgaattt	agaaccaagc	21600
aacttgaacg	cgacacaaat	aaaaaagtca	aggaaaggga	aaatggtgag	catgaagaaa	21660
tgacaagtga	ggaggaagaa	gaggaagatg	aatatgaaga	aggaggttgc	ctatcagata	21720
ttgacgagga	agatttctat	gaggatgggt	acgatgaaga	agagggcgat	gacaatagaa	21780
ctagaaaagaa	gaagaaaaatg	gaagaagatg	aagaggatga	agaagaagaa	tatgacgatg	21840
aagaagatga	agaagaggca	gaaacttgtg	gtgctaattg	tgttattgat	tgtgaagacg	21900
atgcaatcat	ttccccaat	ggacaaaatt	caaaaaggaa	gaaaaatggt	aaaaaaacaa	21960
acattaaaaa	gcggtcacgg	aggaaggggg	agtgtctctg	taacacttta	tcctttgtgg	22020
aaaaatacgt	tggaaattgt	aagagcctag	gtataaagcc	agtaggggtg	ccgcccccat	22080
ccactgaatt	tacttcccta	tttatgaagg	ggaagtgaagc	tgacagctgt	tataatactt	22140
gtcagtcac	cagaggggct	agccgtataa	ggtcactact	caataaatac	tctgttaaa	22200
atgtgatgca	ggtaaacagc	ccttcgagtt	ggaaatgggc	taaccctccc	gaccgacggg	22260
tcgtgctggt	tgataagaaa	actaagggaag	aagttgaggt	taagtttgaa	attgaatgtg	22320
aaaaatccga	gtattttgat	gtcgtatctg	aactccctag	taatattaaa	gtatggttaa	22380
aagagacggc	aaaaataata	aaacatttgg	ctctgattga	agactttctt	ccagctatgt	22440
gtgctgctac	cccaaaaatt	cccctcaatt	tgattaaaaa	tatgacgagc	attttctctg	22500
ttagagatat	tggtggattt	aaaataccag	aagaagtgtc	cagttttatt	cctatagaat	22560
ggaagacatc	tatttctgca	atggggctcc	tctctgtaca	atgtgatcgt	ataatagaag	22620
tgatagattt	aatgataact	aatggcgctc	ttgcgacgtc	atgcttgaac	aacgcattct	22680
tcttgaaaag	aggagtgggtg	cccagagatg	ggagtaaac	gtggctccac	acggaccttg	22740
tgcaactctc	cacctccata	tttagaagta	ttcgcaacag	aggagtgaat	attggcggta	22800
acaacaacac	tggtagcaat	tcttctagtt	cttcttgtgg	agggaataag	ggcgattatg	22860
gagtacgttg	tggtattgagt	ataagcaagc	gtgtgtataa	cctgaaacca	ccacctgcag	22920
cgatgactaa	ttcttcttcc	ccgtcatcct	cgccatgat	ctcattgcct	cagcccacgc	22980
gccagagcat	agatctttcg	ataacgacaa	tcatccaaga	tttctcagaa	gtttctggga	23040
aattgaggct	taattggatta	cagaaaaaca	gtctgtacaa	gagcaaagac	gtgttttaag	23100
atgcaatata	cgactctggc	gcattcaagg	cgctcctaac	atgcacagtc	aacgataaaa	23160
gtagacgtaa	aaggaaaaga	aggactttat	tgcatctctg	agaggggtgtg	gtacgaagaa	23220
acctgatggt	gagtcagggc	aatgacgtca	atgatgccc	ccagttccag	gaagaatgcg	23280
gaataaaaa	tgggggagg	gcttctaggg	tgtataaaa	agcccagcgc	cgaggttcgg	23340
cagtcagttc	cagaagaaga	gtaaggaaac	aaccccagtt	tactatagca	gtctctgacg	23400
aagacgacga	ctgcgaagaa	gaaggcgact	tttcttccga	gttaaatacca	acgcactctc	23460
aactacttct	attccaacaa	cggcaacaag	atagctgcac	agaagacgac	gacgttttag	23520
tgtctgtaga	agaatataac	aacagagtaa	gcggttcttc	caccacagcc	ggagacagag	23580
ttcttgcaaa	ggatcttctc	tctactgtat	ctccgaacga	aaagagggaac	tctgcccgcc	23640
tcgcccagct	caccatatcc	cggcactctc	tttcaacgc	tctatctgca	aaaacaaggt	23700
tgaggagaaa	tggtgctttc	ttctataaag	agcactatgg	actaccacaa	cattgaagat	23760
atggacgatc	tccagcgcgc	cacctacaag	gatcgatgg	agacggaatt	ggtcctcgag	23820
atggctaaga	aggagggaag	gtacgtccga	tcgttggcca	ccatggacga	attggaggta	23880

cctgaagaac	cagccacttg	ctacacttgc	ggctacacct	ttattagacg	cagggcacc	23940
ccaccaaacc	gcaagtcaat	attcagagag	ccttgcgctt	acccagaact	tctccccgat	24000
gcaccatccc	ccgtccgttt	agaagagctt	gtcgcgctgc	cagaaggagc	gagttttttc	24060
acctaccctc	cctacgacga	cggatcttct	acatcgctct	cacaagccga	atgtgaagat	24120
gattatcctc	caccatacga	cccatcagaa	aatccacaga	ggtcccaagt	gtgtgattat	24180
tgtaccacac	gtcaagtcct	cagttctatg	acggatcacg	ccagggccaa	cctcataaaa	24240
aatctgaaga	gggagaagaa	ggcctgggtg	cctggccgct	gcaacaactt	tagctactag	24300
ggttgtacaa	gaagaaaata	ggattgattg	tatcgatgac	gaacttttgc	tgtttttgac	24360
cattcttttg	cttgctgagt	gtacatagct	ttatatattg	caataaaaac	gaccccgat	24420
gtatttttgg	gtattcttta	ttactgccta	tttttatggg	ttaacggaag	aaatttttga	24480
aaagtttttg	agatggagat	gaagggaaaa	agagggcgct	agttcatata	ctgcccataa	24540
gggtggtcgg	tccagaaacg	tctgtctgtt	ccagaaatgt	gtttagagat	ttctggacaa	24600
gtcattttctg	gaaaggggtg	taattttatta	tagttgggtat	attattttctg	gtatacattt	24660
ctagctcctc	tctgtaatct	gacgttgggt	cgacccagcg	gtccaccctt	cgaacttgac	24720
ataaggccta	gtcagcgggtc	caccccttaa	actggagtg	gctgaaaaaa	tttttgaaaa	24780
gttttttgaga	tggagatgaa	gggtaaaaag	aggggtgaaa	ggatagctga	ggccacctgg	24840
gtcagccgtg	tccagaaac	gtctgtctgt	tccagaaatg	tgtttagaga	tttctggaca	24900
agtcattttct	gaaaggggtt	gcaattttatt	aaagttggta	tattattttct	ggtatatatt	24960
tctagccctc	ctctgtaatc	tgacgttggg	tgcacccagc	ggtccaccct	tgcgaacttg	25020
cataaggcct	agtcagcggg	ccacccctta	aactggagtg	agctgaaaaa	tttttgaaaa	25080
gttttttgaga	tggagatgaa	gggtaaaaag	ctagtaatag	aaggttggtg	aactcggcat	25140
aaccagaaaa	catatatcaa	aagggatgga	aacaactatt	ctttttatac	agtagtttta	25200
ttgtaccaaa	aactcagaaa	tctcatcaca	tgacaaatct	gaatgtccca	ttccgtggcc	25260
ctcaataacg	tacgacactg	gacctgggaa	ctccaccta	catgtttttac	atacagaaaa	25320
ttttggctta	gatgcattat	caacaaacac	ttttaccaca	ttggcattca	ttctcttcag	25380
tttgtcacac	atccacaaag	aggccatatt	ctccatagaa	tcagatccat	caacaaatag	25440
atagaatggt	acacccttgg	cgatccagtc	catggcgaca	atagtaaaag	cttggtccaaa	25500
aacactgttt	ccgaagattt	ctttgtcccc	atccatgtat	tgggaaaaag	tacagtatgg	25560
tccacagtc	ttgatctcct	ccaaccagtc	gttcaaaaaa	tccctaataa	agaaagacag	25620
gcaattcgac	ttgatctgga	aaccaatggt	gagttttctt	gtattttctt	cattgttgga	25680
ggggaagaaa	tacgccaatt	gtgtatctgc	ttcaaatgct	ccacacttcc	ctttatggat	25740
caggcgggtc	tccataatct	ccttgccaga	attggggagt	aggccttcga	agaacaggat	25800
attggggcgg	gtaggaactt	tgccagtcac	gtccatggtg	gcaaagagat	cagcagggtc	25860
ttcaaagact	gaggccattt	ttttggttgt	tgctgctgag	agaggagcaa	gtagtagtgt	25920
ggttactagc	aaggagctca	cagtcctata	tatagactaa	cacccccccc	ccccaaaaat	25980
ctccacatc	tttcataccg	gggtcggtca	gaacccagaa	acggtttgac	tcacatctgg	26040
tacattattt	ctgttataca	ttctagccc	ctctgtgtaa	tctgacattg	ggtcgaccca	26100
gcggtccacc	cttcgaactt	gacattcgat	cgactcagcg	gtccatcccc	taaaactggg	26160
tgacccagaa	aaatttttga	aaagtttttg	agatgaagat	gaagggggaa	aaaggttagt	26220
gaatacaggc	agtctctgct	tgctccagaa	atgattctaa	agattttctgt	gcattatttc	26280
tgggtatacat	ttctagcccc	ttcgtgtaat	ctgacattgg	gccgacccag	cgggtccacc	26340
tccgaacttg	acattcgatc	gactcagcgg	ttccacccct	aaactggagt	gaccagaaa	26400
aattttttga	aagtttttga	gatgaagatg	aagggggaaa	aagagggtgc	taggggtgcg	26460
tccagacatc	tctagagtgg	ctagagccca	gaaacatccc	aactagtttc	tgggacattt	26520
tttctcccac	tgacgcagac	tatatagct	aaccactaag	catattttgc	acacatttca	26580
tccaccatca	ctgccgaata	cactcttgtg	ttgtccgca	tagcctaaca	atggccccaa	26640
actccttcca	gaaatttgct	ccagttatca	agactgagaa	gaaggaagaa	gaaagggatg	26700
aacattgacga	ccctttacgg	cagattgatt	ttagagatag	aaagacatta	atctgcctca	26760
ctgcaaaactg	tgtttcgagg	aagagaaaag	ctggatctgc	acatgatcga	gtatacaaa	26820
tactacgcta	cgggaaccca	tacaagtacc	gtcgcaccaa	tagaacacat	cgaggattgg	26880
ccctctcaat	ggtatcaagg	gaagtaggaa	catgcctccc	tctgcgaccc	atgggaagaga	26940
ctgaagaaaa	ccccatcgac	aagtgcggag	tggcgttcc	gtactccaac	tacaatgaag	27000
cgatggcat	gaccacactt	tacaacgacg	aaagatatat	aaagaagtgc	aaaacaattg	27060
aaggaggaac	aagaacgtgg	gtaaagaaga	accgccaaga	atacttcaga	caagctctag	27120
agacattgat	gatgtcccat	tctataaaac	aatattccaa	ttttattttt	ttcaaggagg	27180
atatggagga	aggattttgt	cacaaactcc	acacatttat	taatatggta	caccctaaaa	27240
aggtgtctgt	tttgtgatat	aataaagatt	gatgttttta	taaatgaaac	gtattattta	27300
tacaaaaaaa	tttattatga	tttactcac	tgggttggtc	tttactcac	caatctggcc	27360
tatgtgtgtg	ttcatgctag	cccctcttcc	acccttcac	tccatctcaa	aaatttttag	27420
aaaatttttc	tgggtcgctc	gagtttagag	gggtgaccgc	tgggtcgccc	taatgtcaag	27480
ttccgagggg	ggaccgctgg	gtcgacccaa	tgtagatta	cagagagggt	ctagaaatat	27540

gtaccagaaa	tagtgtacca	gttataagaa	aatgtgaccc	tttccagaaa	tggcttgtcc	27600
agaaatcttt	gaaccatttt	ctggaacaga	cgtttctgga	gcattggtcga	cccagatata	27660
tggtggcggc	tgtcaatagt	agcctatgta	ttcaatgggc	ttttaccctc	ttctccatct	27720
caaaaatttt	tagaaaattt	ttctgggtcg	ctcgagttta	gaggggtggac	cgctggctcg	27780
gcctaattgc	aagtcccgag	ggtggaccgc	tggttcgacc	caatgtcaga	ttacagagag	27840
ggtctagaaa	tatgtaccag	aaatagtgtg	ccagttataa	gaaaatgtga	ccctttccag	27900
aaatggcttg	tccagaaatc	tttgaaccca	tttctggaac	agacgtttct	ggagcatggt	27960
cggtccaggc	agcctcacag	ttacctgttc	acacccttcc	tactagcacc	ctctttcacc	28020
cttcatctcc	atctcaaaaa	tttttaaaaa	atttttctgg	ctcactcgag	tttagagggt	28080
ggaccgctgg	ctcgccctaa	tgtcaagttc	gaagggtggg	ccgctgggta	gggcccattg	28140
cagattacag	agagggtcta	gaaatgtata	ccagaaatag	tgtccgatta	ttaaaaattt	28200
aacttttcaa	catgtggcat	acaccgtatg	aacgtatagc	caaatatgta	ccacagtcaa	28260
aacatttagg	tgagaaagat	aatacacatg	tataaatttc	aaatacttta	tttgcataca	28320
taataaaaaa	tcatacagtt	tcggactcat	ctgtccacac	atcatttatt	aggaaactgg	28380
atctattata	ccctattctt	ctacgtttac	aagatacttg	ttcatcaggc	actgccattc	28440
ttttacaaga	cactattggg	tcaaccaaca	tctctttaac	ttgatcagcg	tcaggctcta	28500
gggaagtacg	ttctctgcac	acaggacacc	gggtaatttc	tccacgttgc	gtactcaaca	28560
ccaatctgtc	ccaacatttt	tcatacaata	aatgtccaca	gcgtgtcgtc	ctgatgatat	28620
tggtactact	gaaatcttct	aaacacactg	gacaattatc	gctgtcgaaa	cgacacattg	28680
cattcttctc	atgttcttct	agagcccttt	taatttcttc	cttatcccca	gaaattatgt	28740
gttcagccca	aggataattt	gaggatgcaa	tataccocaa	gcgacaattc	agtattggcc	28800
aagaacgagg	tgctttattg	acatcgagac	cggtggcagc	aacaaggggt	tcttcttctt	28860
ggagagagaa	atctaaagca	gctttcaata	tggtgggtat	actaggctcg	tatctctctt	28920
gggcaaaactg	tgtaaatgag	gtatcattga	acatcaataa	tgctgcttca	gttacactta	28980
aagtactaga	tttaccatac	aacatattaa	tctttgcccc	atgggctcca	actagtgcac	29040
cagagacatc	tggtttgtga	ttggcgatgg	caatcattag	gggcgttgca	ccatcgctcg	29100
ccttctagca	ttcactatac	ttctatcgaa	atcgatatatt	atattcttgt	atttttttag	29160
aagagcatca	agaaaattgc	acgaggttgc	aacagctgca	gcattggaata	cgttttcatt	29220
attttcatta	ataaaggaaa	aggccaagct	gtttgttgtt	agaatttcaa	tacctactct	29280
tggtgtaaaag	ggggaagata	ttgcagcgca	aaccattaga	ggtgacatgc	ctttcgacga	29340
aggcttcaga	atgttaagtt	tttcttgttt	gactagtgtg	ctagtatggg	aatatttgtt	29400
aatgatacaa	taacacaata	gagtatctcc	cttttcatca	aaataagtgt	gtttgagcca	29460
cggattcttg	gtaagaaaag	cttcaagatc	ttccatgaaa	cacgagactt	ttacacacaa	29520
tctaacatag	cccttaactt	catgagggtc	tggtatcaca	acgtttttat	gtgatacgat	29580
ccttttcaca	ataagattat	cgcacagttt	gtcgggtctt	tttcttacc	taatctccct	29640
caaacagatc	ttgaagtaat	tgtaatcaat	ctcttcaaag	ggcctactgt	taatcatcca	29700
caacagatatt	gcaccacaca	aaccactacc	cactgcttcg	tttattgcaa	catttttttag	29760
aatattgtcc	atcatttctg	cagagaattt	tggtactctg	tcaaatatac	tgtaaagggt	29820
ttgaaaatgc	tggtattgtt	cagccttctc	tataaacaca	tcaacaaact	ctactgccat	29880
ttcagcaaca	gtttttgtag	tagtctcaca	gtactcttct	accatttctt	cttctatcga	29940
ggaaagaaaa	gacttgataa	tcacactcct	gtaagcacta	gccctttctt	cggttggtcca	30000
attcaatcta	gctatagatc	tagcgtccat	gttcatattct	tctctgaagt	cagttgacac	30060
ggatgatggt	gatgtttcta	ggcaagaaaa	aggtctcccg	ataataaaaat	tgccatttga	30120
tatcagtcgt	tttgcccttg	taacacaagg	agattcgtcc	acaaaataact	tgtatccgaa	30180
agatatgtca	aaaggttcaa	gtggtgcaga	ttttttcatt	tcagccacgt	aatcagagggt	30240
gatattgacg	attcttgaaa	agagcctgaa	tctaataaca	ctcgaaacatt	tttcaacgta	30300
gaaaacaata	ccacttcttg	cagaactagt	agactttttc	aggctagcca	aaacaccgtc	30360
caacttcttg	atccttctca	taaccttctc	tcttcttctc	tcagcctgtt	cctttggaagt	30420
aaacttgaat	ccagttctgc	tgtcatcacc	agtgcacaaac	ttgatgccgt	gcgtctcgcg	30480
tctcaaaaat	ccattatcca	tagagaccag	aagagaatat	tttacgaaca	aaaagtcgtc	30540
gtggatgttt	tcgtaaaggc	ctctgaagggt	tttgacagcg	gttgtcaatg	cgttgataaa	30600
agtcattccc	tcgcagatgg	gggaagaatc	agactttgga	ttgtttgtga	taaagaagta	30660
gataaatatct	ctaaactctt	ctttattgtc	taatttcttg	aaactacttg	aaggacacag	30720
aggagaattt	ctggaggtaa	ttatgtcatt	cagaaggggc	aattcccttt	ctgtgaaaac	30780
gtccagaaat	gacatatatg	gttcaatgtt	ttcaagtact	tcttcaagca	cctgacggta	30840
tcgtggagct	gcttcagcca	tggtgatgat	gtctcacata	cgactgttga	gtttatccat	30900
gcgtacgccc	gcttttatac	aaagatcccc	tgtaagaaac	tcctctccgt	tcagttcagg	30960
ataggggtgt	gtcccagttt	tacatccaaa	gttaatatat	ttttttaata	taacaaaaaa	31020
atcgtaccgc	ttattggctg	ctataaaaga	ggtagcacct	gctcacttgg	acatcattaa	31080
ccatcatcaa	tatggaggga	gaacatcaat	atttgaacct	agtcaggagg	atcctagaaa	31140
gaggagtga	gaaggacgat	agaactggaa	caggaaactct	atccattttt	ggaccccaaa	31200

tgagggttctc	tcttcgagac	gacactattc	cagttctcac	taccaagaaa	atcttctgga	31260
gaggagtgtg	ggaagaactc	ttgtggttca	tcaggggcaa	tacagacgcc	aaagaattgg	31320
ccaagaagaa	gatacacatc	tggaacgcta	atgggtcgcg	ggaatttttg	gacagtagag	31380
ggttatacga	tagagcagag	ggagatttgg	gacccgtata	cggattccaa	tgccgtcatt	31440
ttggtgctga	atatgatacc	tggtcttccg	attatactgg	aaagggtatt	gatcaattgg	31500
ccaatatact	aaagaccctg	agagaaaatc	cagatgatat	aaggatgatt	atgacggcat	31560
ggaatcctat	ggatcttcac	cttatggctc	ttcctccatg	ccacatgact	gctcaatttt	31620
atgtggctaa	tggagaattg	tcgtgccagt	tgtatcagcg	aagcggagat	gtcgggttgg	31680
gcgtgccctt	caatattgca	tcatactctc	ttctgactca	tctgatggcc	agtatggtgg	31740
gtctaaaacc	gggagagtgt	atcctcactc	ttggtgacgc	acacatttat	aatacccaca	31800
ttgaggtgtt	aaagaagcag	ttgtgccgcg	tccctagacc	attccctaag	ttgaggtatt	31860
taatggctcc	agaaaaaatt	gaggacttta	ctatcgacat	gttttatctt	gaggggtatc	31920
aaccacacag	tggaaacttg	cagatgaaaa	tggtgttttg	aatcatgtta	aggaatttcc	31980
ttgttactca	tttattccta	gaaatggtgt	aatcgctgtt	gtgggcggag	catatttggtg	32040
tatataagag	cccgtgttag	ctcctcgatt	cagtcacaag	agcgcacaca	cacgcttata	32100
actagctctc	tctctccact	caagatggcc	tttaattttg	aagactctac	aaatctcttt	32160
gccaatatgg	acttgacggc	tggcacaaca	acagacccta	cccgcccaaa	tatcatattc	32220
tttgaaagtc	tactcccca	ctctggtatt	gaggtgatga	agaggcgtct	cgtacggcaa	32280
ggaaagtgtg	ggaattttga	agcaagtggg	ggtgctatgt	cgtatttctg	gctcgaagat	32340
aatgcagaag	atatggagaa	tctcaacagt	ggttcccatg	tcaagacaaa	ctgcttggca	32400
ttattccttc	aagagtttat	cagcaactgg	attgaagaga	ctgatcgaca	tggaacgtac	32460
tgtacttttc	cccaatacat	ggacggtggg	gatgggtcac	gtgggggata	ttttacttcg	32520
ctagccatga	aatggatggc	tagggatgtg	actttctttg	tgtttgttga	taggaataat	32580
actgtagaaa	atgcggcatc	catatggatg	taccaaaaa	tactagcaat	tggtgcaaa	32640
gtagtaaagg	tgattgttga	caatgcatca	aacccaatgt	tttctgtatg	taatgcgtgt	32700
aggtgcaagt	accaggccc	agtgtcatat	gttattgaag	gccatggagt	gggtcattct	32760
gatttgacat	gtgatgagat	ttctggattc	tttgtataat	aaaaccccat	aagaaacaat	32820
aatctttttt	attcaacacc	catgatttta	gtttttatag	tatataaaat	caagaataaa	32880
gtatgtagat	gtctactttt	cgatagctcc	ttcacctata	tagaggaggt	ggataaaaac	32940
agaagatatg	gattctaata	cttctatttt	accgccaaag	aaacggccag	ggttaaatct	33000
gttacaggtt	ttagggatta	taataacggt	agcattaata	gcttccgttt	catcctttat	33060
gttttatagg	gtaggtaaac	gcaaattatta	cccttcttca	tcctcttctc	cagaattatc	33120
tgatgtagat	aatggggtag	aaggaggagg	aggaacaaca	acgacaccaa	ctcaaccttc	33180
acctgacggt	ggagatggat	acgtagatct	ttctcctcaa	aagaaggctg	aactaagaac	33240
tagagttgca	aacgtcatct	ttcaagaagt	gtcaaaggat	caaggagtgg	ccttttagacg	33300
ggcaatgaat	gattcaactg	ataagataat	ggaagaaact	gaggcgagaa	tcaataaactt	33360
ttcagagcca	ttcagagaag	caaccgtaga	acgtgaagtg	tttaaggatg	acacagacaa	33420
aaactttatc	ctttcaactc	tagatttaac	agaggaacaa	tttaaggaca	tttgttatggc	33480
tgaagtgaat	aatcaattag	aaaattttga	ctatgaagac	atgacccgtc	tcattctttga	33540
taacatccca	gagactgatt	atztatggac	aactcatctc	gatccgaaaa	aatatgacac	33600
gtactctgaa	aaggtattag	ggttctcaga	tataaatagt	atagaaagaa	tatcctctac	33660
atctttataaa	ggtaaaaaat	atgaggtaac	tactggaat	gtagctgtcc	tcgttgattt	33720
tgaattctgaa	acaataaaa	agaaggcagg	aaatagtctc	atccgtaatg	tcgagtttat	33780
tggttgaggac	gaacagacct	acaaatcttt	cttccctgca	ttcaatcaag	ttttcttctc	33840
ctttaaagta	aataaggaga	aaagggaagt	tactgtatcc	atcaataatg	gatgtgtagg	33900
tatagtggcc	aatattactc	ctctaactac	gccagttgga	gcagcttccg	gacactacat	33960
catggcact	agcacagcaa	aggaaaagac	ctatctatct	gtaatagaca	agtacgatac	34020
cactgaattt	gtttgtggtc	tgagttaacaa	gtcaactcct	ctcatggctc	taaatattct	34080
ctttatgagt	gatactgttt	tcccttcatt	tgacgaagca	gaaagacctc	tgacggatgc	34140
caaggcagta	gaaattttag	gtaaaagact	aggtgtagga	agatacacaa	acgccaacat	34200
cagaaatact	cagttagatg	gaaggggtta	ttttggataa	gatagaaaca	attgcaaaaa	34260
gggcctctcc	tagttatggg	tctattgatg	tggttacggc	tattttgcgc	cgtaatttca	34320
tggaaaaaat	taggggtaaa	ataaatgaag	aaaccacaa	ggagaagatt	atgggcacaa	34380
aggaagaaa	agaggacact	ataagaagta	taagtggctaa	tggtatcaaa	gagaatactg	34440
ttaaagaaaa	tgtaaccgaa	aaaatttagag	caatgacaga	taaggaatta	aatgacaata	34500
gggaattttat	gcattgattt	ggaaaaattt	caactggaga	tggaaggaacc	ttccatctct	34560
ttgaagatac	accgggtttt	gaaagtgcct	taaaggcaga	atataaaaa	gttccaggag	34620
caactactcc	aaaatacgtat	tctatgaaca	gtttacgtat	cgatgcgatt	aatggaaaaa	34680
tcgaagaggt	ttataatcct	tcacctatca	tggttattag	agaatacggc	accattcgca	34740
ggggcaggta	cgaagaaaat	gcaggttcga	aagaattggg	ttttatgacc	aagattgaaa	34800
aaagacccaa	taatgtagct	gaaaatctca	ttattagagt	tgccaaccag	cagtataatg	34860

ttatgaggat	ggtgtttttt	atagactacg	aaacaaagaa	gggggtgtcc	aaggaggaaa	34920
tgtttatacc	atataatggt	cagaaaacaa	aggctcttaa	agggcgtagt	acttactttt	34980
cattcgtaag	gaaaattcct	gatgaaccag	aaggagtagt	cataatacac	gcactagggg	35040
tttattgagg	aaataataat	aataataata	ataatggcat	tacaggaaaa	ggatataact	35100
atagggaaag	tttctgctgc	cctacgagag	ttgatgtact	caccacacac	tatgcagcat	35160
cacgataagc	taaacacatt	cctggacaga	aatgttgaat	catcttcaga	agagaaaaata	35220
agacaaattg	tggataaaat	acgatcccaa	acaacatctg	acatatctga	aacagtcaat	35280
aatgtcacaa	ctaattgggac	tgcatttttcc	cttttcgaag	ataccttaga	aggtatgggtg	35340
aaaaaaaaata	taggtgataa	ccttcagagt	ggggacttta	ttgatggccg	taaaaagctc	35400
aatgacatga	agagtctagc	tactggagcc	atcttatcta	gacagcgaga	ttttgttgca	35460
gaaagtataa	caggaacaaa	ggactggctc	aaggctataa	tgggttgtgg	tattataagg	35520
tatactgtat	ttgtcaataa	ccttgcaaga	tcaacactcg	ataatgatga	tgacaaggca	35580
gcaacctatt	ataacacccc	tatatatggc	gggtattgta	aaatggctat	aaaggactat	35640
gaaataccag	attcgtagac	caaggtcgaa	gcggaacata	cagttgaagg	aagaaagatg	35700
acctttaata	taaaatggag	aggcgatacc	ataaacaacc	taataacaat	catcccttca	35760
gtgacagggt	atcttgcttc	catctctgaa	gacgcagatg	tcgaggcgcc	attactttta	35820
aactgcaaca	actgttttat	agaggcagat	atgagttagc	tctacatgga	tgagaaaaaa	35880
acagaggcat	catttaccct	caacttaccg	gaaatcgaag	gagctgatgc	gaatgcagtc	35940
tatgaaatat	gtatagtagt	agtttgatgg	aggacaaata	aaaaattgga	tacatttcat	36000
gtagttttat	ttaaactctta	atcttatata	catttcatgt	acatttcatt	aaaaatcact	36060
atcacagcta	tagtcagaaa	actgggtgct	gagctgatta	atcctcctgg	ccatgacggc	36120
agatgtcctc	cttgaaagtgt	tgcgtctgcc	gtctctgacc	ctatcaacca	tcctactctga	36180
acctgggttc	aaattaacat	ttggtgcctc	aaaggggggtg	tgaagagggtg	aagaataagg	36240
agaagggtaca	ggactagcat	catagtccct	cacaatccct	gttcttcccc	ttctaacagg	36300
tgtggatgcg	ggagggggca	tattagccgt	gtctagcaac	tccaacatct	gtgcctggta	36360
ttcaagtgcg	ttagacacaa	acctctcagt	gtcttctgct	tcattgtcat	tattaggaat	36420
aacaccataa	tcttccatga	ggcgctcacg	ccttctcgcc	tgatgcctct	caaagagcog	36480
ttttcttttc	tcagggtcat	cagagaagag	tttgtgaacc	cagcatgggt	tcttagcctc	36540
ccctacagcc	aagtcaggca	caaaaactgg	cttatttttt	acattcttat	agggagaag	36600
agggggcaac	accttgacag	gagcaggagt	ggggggtaaa	ggagtgggtat	attcaccctt	36660
atatgaacta	ggtttatgga	actcctggca	aggtattttg	ccctccctgt	cctcttttag	36720
aggggcggcc	ttacagcctc	caaaaacett	cttagctatt	ttagtatgca	ttttctcaat	36780
ccttttcata	ccttcagaca	acatcatgaa	ccattcttta	tcaaactgaa	aacccctcaa	36840
gaaaaactta	taaaaggact	tttcattctt	tgcaagaaga	catcccgcca	ttagccta	36900
tgcaactcaa	acatccccct	tctccccctc	ccataacttt	atttcaggaa	acctattgag	36960
tgctgagcag	gtactttcta	tagatgctct	cagagcatca	ttttccaagg	ttgatgagct	37020
ataatcaacc	ttcttacagc	tggcacactt	cttaaccata	tccttttatt	ccattctttt	37080
cttatgtttc	ttcatcattt	ctgatgaaga	ggagtgtttg	attgatgtct	taaaagcagat	37140
attcttccct	tttattgtct	ttcttttctt	gtgacgata	caccctccct	tcttgctgct	37200
gctgctgccc	ctgctgctgc	tgtgtgatgt	gtcagttccg	gccattttta	atgactctaa	37260
aaaacaaggg	ggtaacacat	gtaataaaaat	gcattataat	gtaacattct	tttattttta	37320
taaaaggcag	tagaaaattt	gtacatcatt	tacaggcagg	catatacaaa	aacacatcac	37380
tttatagatt	gggcagggaa	gatatttgat	cagcttccga	agggtcagta	tcaaaccctat	37440
tcaacaatga	actgtgtaca	tgtgcaattt	cgatcatcaa	cctagcatta	ttagaggctg	37500
acttaatgag	acaggaacca	attagctgct	tattagcact	ccactttgaa	gcgtacttac	37560
gcttcacctt	caattttctta	gagcacaac	tcttaccat	tacatcattg	acagttgtat	37620
ggaacagggt	acctctagat	atacaatctc	caagcttatt	ctcagtttcc	tgtctccttt	37680
ctctaagagc	agaaagttca	gccagatacc	tattggcaat	ctcattatac	tcctcatcaa	37740
tattatcgag	tctacacata	tacctcttga	catctacaca	cagtctctta	ctctcagagg	37800
aaatattaga	catcctgtcc	togtcacgcc	cataccttgg	cctctttcta	ctgttactac	37860
tactactact	acttcctgag	gaagtagagg	gctgattcac	tgatgacgaa	gaaggtatta	37920
cttcttcgct	atcagagtct	gaatcaacat	taaaacaaag	gttaacatta	atgttaagag	37980
aggttgaggt	aggggcggat	gaggtagggg	cggatgacgt	agaagcacta	ggcatattca	38040
agtcaggcaa	agacattgtg	tcacatcatc	catcatcatc	attaacatca	taaaagtgaac	38100
tactaatacc	actaatagta	tcaaaggtaa	tattgttact	tctaggggga	acagtcaagt	38160
caggaagtga	ctctgagtea	tcattatcat	ctacctctgg	ggggacaaac	actacagtat	38220
cctcatcatc	ttcaatgaaa	ttgaaggatg	tcacaccaga	caacaaacta	ggggagacat	38280
tattatcacc	ttcatcatca	tcacatcatc	tactcaagca	ttgagtaatg	tcaatagtat	38340
tgggagctac	atacttaagc	catttaccctt	catattcata	aggggggata	tccttgaatt	38400
taccatttaa	agctgcctca	tgattatctt	caaaggctct	ttcactgaat	gcatgcattt	38460
cctgtctaga	tagagtttca	agcatgaaac	tcaacttggg	ataagccctc	tccatataaa	38520

WO 01/38351

12

PCT/US00/28888

catcctcatt	ctcattaaca	cctttggcat	tggggacaaa	gaaagtaggt	gctagaatat	38580
ttgacatctt	attaagggtg	tccttttacac	cttcatcatt	aacatagtcg	gcagacaaca	38640
aaaatgagtg	aaaagaggca	tactttgtaca	cctcttcata	cacttttacgg	aagaaccaca	38700
caggggggat	ataggcccta	cacatgggac	ataacagggtg	atcttcattc	accatatgct	38760
cttcaaatct	ccaggtgaaa	agacactcca	ggtggacact	tttccctgcaa	cacggcatcc	38820
ctccatttga	ccagttatag	caatccgtgt	cogaatcgta	atcacccata	cagatgggac	38880
aggtaggcac	ctcatcaacg	gcgacggttt	cagccatctt	gttgttgttg	ttgtttactt	38940
taaagaggga	gaaaagccaa	cgctatagca	actcatttct	ttacttttaa	caggttcatt	39000
acagtcacct	taaaaggga	tatgtgcttc	atggcaagg	catattctgt	gattgaagg	39060
cgattttgat	agaaatcagt	gtctgttccg	ctttgcaaca	acatcaaaaa	gtgcatgttt	39120
gacatcaaat	taaaggcagg	gccttccatg	ttagaacaag	caacaacatc	ttccagtata	39180
ctggaagcca	ttacactgaa	actgtacgtc	attgaatagg	cacctagaca	gtgtccccta	39240
aagaactcgg	gaggtgtgtg	aatgtactgc	ctgattagtt	cttccgtgtt	cctaaaatat	39300
tcctcgttgt	aacaataacc	tagagggtat	gagagtcoca	agtcaatcat	tttaggttgg	39360
ccagatatgc	tatctataac	aaagttatca	gcctttatat	cgacattaat	aattccctca	39420
tttgcaactc	ttgtaattac	attacaagtc	tcagccaaaa	caaaaggcat	gttcaccttt	39480
agtttttcaa	acatacttct	tacaaccttt	cgggcaatga	aagagtctgt	tattgggtgcg	39540
actgtatcaa	ccccctttac	ctctggtaaa	cgctccatcc	tttgaagcat	gacaacccca	39600
tacaacaagc	gtgcgtccct	tgactcataa	tactccccac	catctaaact	atccttgaca	39660
cacctaata	catcatcaat	actgatacct	gcaaatggca	tctctaaaca	taacccctga	39720
ggacacacac	cagtcacccc	aactaccccc	ttaatattct	ttctgtatac	agaaaggcca	39780
ttaacacact	caaaaacaaa	ctcttcatag	tcataatcct	tcataaaact	gacaacagtc	39840
cccctattca	caatgtaata	gaccccatat	ttaatctcaa	aaaaggctct	ccagggacag	39900
atgaaagtgt	taaagtcctt	cacatcttca	ggcacttoga	ccaaatcaga	caaacaacca	39960
cgagcagctc	tacggaagtt	ttcccagttc	tcacactggg	cagaaacttc	cttaattagc	40020
gtgctagtct	ttgtgtctaat	tcgctcacag	gcacttgaag	gcaatttgaa	cacctctaca	40080
gcatatata	ttatgtccat	aatctgggag	gacggaatat	tacgactaga	caactcctta	40140
agaatgtgtc	tagtgtcaaa	aaatctcatc	tcttcatcgc	aagaagaggt	gggccaatag	40200
aagccaactt	tagggtccct	tccatagaaa	caggatatcat	ggacataact	gctagactta	40260
acaattgacc	tatggccaat	tctcaactgt	tgacagggtat	cataattgtc	ctcaagacta	40320
caactcttct	caatgccttt	aaagtagtaa	ggggaatcac	gttcaatcca	ctcattatac	40380
tgccctttac	actgcctaata	caccttggac	gcacacacgt	ttgaaaagg	gcgagtagtg	40440
ggaggagagg	gagcaacttc	agacttcttg	aaacgtttag	ggggaggggg	cgtagtgcta	40500
aggttgatga	cgctatccca	agtcttgata	gtgatgttgt	tggtgttgtt	gttgtaagcg	40560
ggaacagttg	tggcttcgga	ggttgaacca	gacgcgtagt	tggtgttgtt	atattcttcc	40620
acaggccggt	ttttattccc	ctgatggtaa	acatactgct	gggtgtgtgt	gtctccacca	40680
gtattgatgg	tagtagtaata	tacagtgggt	cccccatgtg	cctatggaag	aaacaacaac	40740
aaatacacga	ttaattcaac	gcccggcaca	ctaataaagt	gatgtgtttt	ttatataaaa	40800
aatagcccat	aatatttaac	gttacgttaa	tcaaaaaaaa	caacaacaat	aatagtatat	40860
tatttggtact	tacttttgat	gggtttggag	cggagactga	agtaaatcca	gaaaacaaca	40920
gttttactgt	tgggcactac	actgcactgg	tgacgtagta	gtagaagtag	tagtggtaat	40980
agtcactat	tgtcactgca	ctgcactagt	agtagtagta	gtagtagtaa	gagcgttgga	41040
taagctcggg	ttggctgcta	gagctggagc	tggaaactgga	gcttgtatgc	ttggctgtgc	41100
ctctgctgag	actgatgctt	ctggcctgcc	ccggctcgct	tatatacaag	ttgtcccccc	41160
tcacccccac	ttccagaaat	ttgccgtcga	acgccagttc	tccaacagag	tggggtccag	41220
atctggctac	gggttcaact	ctggccatac	cctgttttagg	ggtctaaatc	cacccccggc	41280
actaaatggg	gaggaccta	gggtgttata	acaacatgta	agcgttgggt	aagaagatct	41340
ggattctgat	gacccacctt	gtccttctta	tcctatcctt	atccttgtcc	cctgtctacc	41400
accacctcac	cccatacttc	tcaccccatc	tcacctatac	ccccatctca	cccatcacct	41460
ctatattccc	ccaccttatt	cactcgctcc	agtttcaaca	ccctgttctt	gccgagccaa	41520
cccataacca	gatctggacc	ccagttcttc	cttttatccc	taaccggcac	catttatgcc	41580
cccaggcgct	agcgggtgat	ataaggcggc	cggcgagggc	cagaagcatc	agttctctgc	41640
aagccagcag	aagagcaaca	caacaagcac	tctctctcct	tctacctaga	agagacctgc	41700
caatactcaa	gctacaagaa	tggcctctcc	agcccccgcc	gcaccaagtc	cttacacct	41760
gttggaactc	aagttactta	gttctgagga	actaaaggaa	ctaacttcat	acgtctcgac	41820
tagctctcgc	cgtctcgata	tgaagaaaca	cttgctccat	ctattcgagg	agcacgagaa	41880
gatcttccaa	ttcatacaag	gtaagcacia	gttctcacta	tacactttgg	actttgaaat	41940
tttcttggtt	atgtcgaata	ttttgttggt	tgaagtgaia	aatattctaa	gtccaattcc	42000
tttactcttt	gacagaaatc	tccaaccagt	acggagacta	tggatgtttc	acaatggccc	42060
cgctcaccct	gaacgctgca	gccgatctct	tggataaggt	gatgtccgga	cctctatctc	42120
ccgaaggcgc	ccaaacttgc	tccccggctg	cttgtgttgg	tgcgaaggtt	gtgaaggcac	42180



tggtgagctt	ttgccagaag	acccgcttca	ccaccaacat	tgtgatgaga	gaagttaaag	42240
ccatggagtt	ccaaggagac	gattttaact	actctgcctt	gtgtgcaagt	atgccccaac	42300
gccccgtgac	tgagaggcag	atgttcgccc	ttatgaagag	tgaggacgaa	gaaatgggag	42360
tgtctgcaaa	cttctctcca	gtctctgatg	acgtcatcaa	cccttcaagc	ctccccctctg	42420
gacaagaagt	cgactcatca	acttccgctc	aaatttctgg	tatgtttcaa	aacgtgtgga	42480
gtttgcttga	agagtgtggt	agtggctcta	atagtaatag	ttccccctgtc	tctaggacag	42540
tcttagtttg	cacctgtttt	ataatccaag	tgttcaagtt	tttggtgact	aaagtgtcta	42600
atgtgaacgt	acttaaccag	ttgtttggac	atgtttgtttt	tggtatcactt	gatgtggctc	42660
caagtaataa	taatagtgtc	ccatcaactg	ttgttaacaa	caacaacaaa	ccctcgacct	42720
ctaataatag	taacaacatc	agtaacaagc	gtgttggtgg	tagtaataac	agtggcggcg	42780
gaagatcaaa	gaaagttaca	gccacagcca	aaaatccctt	taataatgta	gatggggaca	42840
atcatggcat	gtttgcccgt	gccccgtgtg	atgttaattt	ggatgacttt	gtttttcccc	42900
aagttgaaac	tottacaagt	aagagcacca	tccctaaaga	agaggtaaat	gtagatgaag	42960
atgtgagtaa	aatgtgccgt	aaaactgccc	ttacccccct	agaaattcat	acctttaatg	43020
tgttcatctc	tgagattaac	ccctccaaat	atgaccgttc	aatgttttgc	aagggattttt	43080
tgactgcatg	ggataagttt	gtagaggggg	atactgtctg	cgttaaacgc	ttccgttaact	43140
atatectcac	tcgctcaaac	tatgcctcag	ccgccagggc	cgtgtatgaa	gcgtcaatta	43200
aggggactgt	ttattataat	gacaagtcaa	agttttctgtt	ccacgataat	gttaaccctg	43260
atctggacaa	gagctggggg	aacaagaatg	ggaagaaacc	tagactccca	gctaacttga	43320
tggcattcat	gggtattgac	attgtaaagg	tgtgcgctaa	ggggattcaa	aagtatatgt	43380
ttgcaaagca	attccaacat	ccggaagtgg	aagaacttgt	gcctcctatg	gctgtatacg	43440
caaaggttgc	cgcaggattg	aagtcggggg	ctttgtttga	tgactgggac	ctgcctgaat	43500
acgaaaattg	tcagttttatc	aagtatgaca	cagaagggtg	caaaaagcac	agtgaattat	43560
acgccaaaaca	acttctccgc	acaggactta	atcaatacaa	taaactggaa	gagggacaga	43620
gtgcattccc	atltgcaaat	attgtgacgg	taacatccgc	ctctagtgat	gatattcacg	43680
gtgacacaat	cattgaattg	atgtacaaga	caaaggatgg	cgtaaaagga	gtctcaaaaa	43740
ttgaggacga	aaacatcatc	aaggtgaatc	cagcagaaga	aaagaagaat	aatagagtac	43800
aagccgagaa	gacctgttat	tttgagattg	attccgatga	tgaggtgtgt	gagagaacag	43860
aggaagaatt	cttcaggcct	acatctgttg	ttgttgcccc	gacaacaccc	ctcgtacctt	43920
ctaattgtgga	ggaagaggaa	gaggaagaag	agcagatgga	agaagaggag	gaagagggaag	43980
tagaaaagga	agaaggatct	gataaggaag	atgacggaga	cgaccagca	caggaagaaa	44040
tgaggagga	gaaggaaaga	gaacaacaac	aacagccaga	agaagaaagc	aatggtaatg	44100
agaaccaaga	agaagaacaa	caacaacaac	aacaaccaga	aagagaagag	gagaataagg	44160
atgcagatag	tgacagcgac	agtgatagca	gcagcagcag	tagtagcagc	agtagcagca	44220
gtagtagtag	tagcagcagt	agtagcagca	gcagtagtag	cagtgaanaa	gaagctgaaa	44280
agaagaaaaga	agaggaagta	cctgccaaaga	ttcagaagag	aaagaggcta	agtgaaaaggc	44340
catcagaagc	tgcttccctc	cccaagagaa	tgagagtaga	agaagaacaa	caacaacaac	44400
tatcaccatc	attggacata	ctccagactg	cagttgatga	gatgatggaa	gaaattcctg	44460
cgcttgagcc	tatcgttgct	acaacctcac	ccaaggcagc	gacacttgca	ctcaagacag	44520
gatttagtta	ctcttcattc	gtaagaggag	atgaccttcc	agtagctggt	aatacttccc	44580
ctactgaacc	agcagctgtg	cccgtgctg	ccacttgcac	ttccgatgtt	ggaaatgact	44640
ttttggacat	gttggacggg	ttacctggcg	atatagtaat	gcaacctggc	gaatgcgacg	44700
tgaccgcaaa	attcttttgag	ggcatcaccc	taccagatgg	tactgataat	gaatgcacag	44760
gtttcgatga	tcttctttaa	gccaccgaga	ctgataacat	tataaccacc	acatgcttta	44820
cctccccgat	tcacctttct	agcaactcag	cccccagaaa	ggatattgat	aattgcagtt	44880
ctattaagag	gtctagggca	ggttcacttt	ttgacactga	tgatgatagt	gaaacaaatg	44940
aggttgaaaa	ggaagcccct	aaacgtaaaga	agcacttgaa	aaagaggcgt	aacaagtccc	45000
accgtggttc	ctctggttct	gcttcttctt	ctcattgtat	gagtagtgat	gaagaatcag	45060
aggatgaaag	ggatatgaaa	tcaacatcaa	aggttcacia	gtcaccaaaa	gctcatgtta	45120
aacattcccc	taaataatgat	gctgtaaaata	gtgatgtaaa	taactcatac	aacaatgtta	45180
atagtacaac	atgcatgtcg	tcacagata	gtgatgcaga	agcacagcct	aaaagccata	45240
ataaaaagcca	ctctcgtaaa	cactcttctt	ctcccaacaag	tgataagaaa	cagaaccaac	45300
aatgctcaat	caataactcaa	aatgtcaaga	agactgttgt	acagctctcca	cctagtttta	45360
gaagtttttag	tcctaagaaa	gatgagcttg	gtgattttctt	gtcacgcaag	cacacaaagc	45420
cagttaggcc	ctataacaag	aagcgtgata	atgttaacac	cactaataat	gtagtacaga	45480
ggtctgctctg	accgactcaa	atgatactca	atcaatgtac	aataataatc	ttagtactta	45540
acaagaagaa	ctatatTTTT	ataatatTTT	tacatgtctt	aataacaaa	aaaataaaaag	45600
aaaaccaatg	tattatatgt	ttaaaatcaa	ccccatttgc	atgattaaac	taactatagt	45660
gtgtaaggaa	agaaaaaaaa	catgattatt	tctgccatta	aaacaacaac	aaaaattcta	45720
agcttctcct	ttctttctgt	gtctttgcag	attcgggagt	ctttctaaaa	ccacagcaat	45780
acaaaaaaca	acacaaattc	tctactgtct	tccctcacca	atccccctat	cgccttatg	45840

ctccttctcc	tctctataact	tcctccccat	tcacttctga	ctatcaattt	tctgattcat	45900
ttcaacaaca	atgccacca	tacaattact	acccacacc	ccctccatct	atgtcctcga	45960
cgacctctcc	atcgtctttg	tgggatgatg	acgatgatga	tgacgaagaa	gacgaaaaag	46020
atgtcaagca	agaagtctcg	aaccgtcccc	ccattttttc	gtacatggaa	actgtatcct	46080
ttagtataa	cgatgaggac	gataacaagg	gagaagaaga	atgttttggg	tcaaactttg	46140
atatgttttg	tgattcagat	aacatgccat	caacttctac	tgcccccttc	cctcctccct	46200
ctacaacaac	accacttctc	actcctcgat	ccatcatgga	tactgattcg	gatgaatgtg	46260
acgaagaagg	agcagcagca	gcacagcac	cgtctattgc	cgctcttctc	tctatccctg	46320
tcgggatctc	tgaagctgaa	ttgaaaaaaa	tggaaaagaa	aaagaggaag	gaaattaaga	46380
aactcaaaaa	gatgatgaaa	gacccctctc	ctcacctata	tgtaggagga	gaacctcctg	46440
tcgcagcaga	ttataaaaaca	agggcaaaaca	ttccctttta	taaagttgac	cctagtatcg	46500
atatgtgcgg	tgctgcccc	cctcaatttt	gcgctgaatt	gcccacccca	tccatagatg	46560
tgtatacttc	ttcctatgta	tttccctctc	ccacacctgc	catgcataat	aagaaagggg	46620
ccaagaaatg	tcaattcctt	aaggggagaa	aggctttgag	gaaatggatt	cacgagaatg	46680
tatgcattgc	ccctcccggt	aaaaggggag	gtgtattttt	ggctcacttg	gaacaaagat	46740
tcttggtcga	acatggagat	gaatacaagg	tcocaaaggat	gtttgtttca	agagtattga	46800
acaaagcttt	ccccaatctg	atgtctcggt	cagacacact	gtgcagtgat	atgacattct	46860
atactaacct	ttgttggata	gttaatggag	ttgtcgtatg	ctttgataaa	gatgatggag	46920
gaatacatgg	cgatgcgtca	gagtatgcaa	caggagaaaa	ttttgatact	gtagtgttcc	46980
acaagaggga	agaacaaaag	accaatggga	gtgccagtaa	gaagaggcgt	ctcacgcctg	47040
acactagtaa	tatgggaaca	agcactgatg	tgaagaat	ccaaacgatg	ggaacaaaat	47100
ctgatattga	agaattccaa	tcaatgggaa	ccaataccaa	ccccatagag	acttcatacg	47160
tgggtgtgaa	taccaaccca	cttcccaacc	ctcccccaag	atttgtaatt	actcctttta	47220
cgaatgatgt	accagaattg	gacatgatgt	ggctttattc	gccttcaga	ggaggtggaa	47280
attctagaat	gagtgcacaa	acaggaacat	ctcccctgtc	taacacccca	attcctacct	47340
gcttcacagg	agggtgcgaat	gtagtatgct	ctaattggatt	tgtccctccc	acgttcccct	47400
tagaatgtga	cgaagatgat	ccaagtattc	ccaattctta	caattacgaa	gaggataaag	47460
tctttcatcc	attttatgag	tatatggcca	aatatctatc	ccctcttgtt	ccatcatata	47520
acaagggaca	gacttgtaat	gttgtccagg	agtggttcaa	gggatccttc	tctcttgcaa	47580
agcgtagagg	aacagtcctc	aaattctgta	gtaacatttc	ccacgccttc	ttttgtaata	47640
tggatgtatg	tactgccatg	tgcaaatggg	cgaagactgt	aattagacat	ggacaatatt	47700
gtaaatgtga	tatcgtaaag	aggtcatgta	catccatgct	cgcatatcac	tacattgttt	47760
gcagagacgc	ttcatgtgat	gttcccaagt	gcagggaaaag	ggttcgcaac	gacatggatg	47820
actgattgat	tggttgatat	gtgacatttt	ctgtatatgt	tgtaataaag	ataccaataa	47880
actaatgttt	tatatatgat	tctatttttt	taaaaacctt	taaaaatata	catataaaat	47940
gatgtatttt	tgaactaca	ctctggcaga	atcagaccag	acccttgacc	taaagcagac	48000
cacaggggag	tcttagagag	gggtgtgaat	ctggctaggg	gttcacccct	caatgttaac	48060
cacacggcaa	taaaaacacc	acttcctaga	aagggaggga	aggggtgact	tggtgatatc	48120
ataactgggg	aatttctctc	ccagatatct	ggctgtacac	gtgtgagcgc	ttctgggcgc	48180
gacaagaaaa	aaattagtg	tatcataact	ggggaatttc	ctctccagat	atctggctgt	48240
acacgtgtga	gcgcttctgg	gcgcgacaag	aaaaaaatta	gtgatatcat	aactggggaa	48300
tttctctccc	agatatctgg	gtaagaaaaa	aaattagtga	tatcataact	ggggaatttc	48360
ctctccagat	atctgggtaa	gaaaaaaaat	tagtgatatc	ataactgggg	aatttctctc	48420
ccagatatct	ggctgtacac	atgtgagtgt	tctaattctc	ttttttatat	agaaaaatata	48480
actgaattag	ctctctaaac	ttttctctcat	ttcttactcc	tctttgtttg	tatgtcctac	48540
taccagtgtg	catataagaa	ctggagagat	gggagtttta	aataacaacc	cttcttaatc	48600
cctttcatgt	ttatatatta	taaatataat	caccatggat	atttcaataa	agacattatt	48660
tttagtagtc	gggtaccttt	ttctgaccac	ttgtgcatct	tgcaagtcaa	ctcaaatcgt	48720
ctggaacctc	atggtagctt	catttggttg	atttctagga	cataaactac	ttaaaaacat	48780
tacacctgtc	aatctggatc	ttgtcggaag	aagcttcgta	ttctctgcaa	gtttaaccat	48840
ctcagaagaa	gcccgtttat	tgaggattgg	aaacgtatta	agggactata	atggcaacaa	48900
ttttgaagag	tatgaagaag	aagaagatag	tggtattgaa	gaataataat	aataataatg	48960
gtttgaataa	aatatagaga	tacatttata	ttgttttatt	tgcatatat	ataaaaaagc	49020
actacaaaat	ttgtacacat	acattgagag	aaaaaattga	tacaatttct	tccttttttt	49080
tactgggtatc	tgatttcttg	atattcgaga	gagttagtag	tagcagaaga	agtagcacca	49140
actccggcag	cagaagttgt	agggataggg	gcggtggcac	cggtggcagc	ggcagcggca	49200
gcagcttcac	ggcggtgttg	ttgttccagg	ataatctttg	cgcttttatt	catttcaagt	49260
ttgctgtgaa	agatgcctta	gttgtatttg	accaccttat	tcctatatct	tcctagtctc	49320
gtatcgcgct	tctcaaaaaa	cttggccatt	ttccacacac	gttgacgggc	tgcccatgca	49380
gtcatagcca	taaaaatcaa	ttggtcagag	gtgatattat	gggtgttttag	aaggtccaca	49440
gcgtaggcca	tacagctagc	ggcacgttga	atattggagt	acttgctgtt	tcttggttta	49500

acggtatgga	tcaccttact	tccatcatte	aagaattctg	cctcttctct	gtgcagaaga	49560
gatgtgcttc	tgctaactgt	cccgtcgtcg	aggacaaaag	caggtccagg	tatgggtgtt	49620
tcgcaatctc	cgtaagaagt	gtcagcaacc	gtgaagtggg	aaggatatag	tccattatct	49680
ccaactccag	aagtgaagtgc	agcgtataaa	acttctccgg	tgatttggcg	agtgaactga	49740
ttacttagcc	ctgcgtcaat	ttcagacata	tttcttcttt	cgggtgtggg	aaggaagggc	49800
ctaggtgggg	cgagagtgtc	gttggttctt	tcagccgtag	tagtagcaag	agatgtgacc	49860
caagaatatt	tcttgttctt	agactgtctt	ccagaaccgt	cgtcgataac	aagagatggg	49920
tcgtctcgt	cactgctacg	gttgccattc	tggttcttct	tcctctttga	atccttcttc	49980
ttagacttgg	acttgttctt	ggggccaggga	acaggtataa	catccacatt	gggtgtagaa	50040
gacatcctcg	ttccagtaga	ttcagtcaga	gatgtaagtg	ctggcacggg	agggatggcc	50100
gtaacagaga	taccttgtgg	tactcggaac	attgcagaag	aagacgtaga	attagagtgt	50160
atacttgtaa	cagcagaagc	agaagtccga	gaggcgatgg	cagcagctct	ggcggcagca	50220
atcatagggg	ctgtagtggg	gcaaatcggc	agagtgttgg	atgaagtcgt	ggcagcagag	50280
gtagagttag	tgtgagacat	cttgatttct	tggttgggtg	gttgggtgtg	gactattgtg	50340
ttctcagaga	ggactttagt	tcgactgtgc	tctaccaga	tccacagccc	ctttcatatc	50400
cgattttcgg	gggttggccg	atactaata	actcctccac	gccctcttct	tgcgccatat	50460
ctataagggg	cgaaagtgtg	ataccttgag	ttgatccggg	tataaccgcc	attgttgggt	50520
ctttctgaga	atcgagagtt	gcgatcacgg	catggcctcg	tgggggaatt	ttcattgacg	50580
taagcagtaa	cgtcagaaga	attgttggag	cggggttggg	taaagtgcct	ctgggtgttg	50640
tggcggtgtt	gggtgtggcg	gtggtgttga	gcagcagcag	gggcgtcgta	gtgttcgtat	50700
tgctcacaat	aatagtcctt	gggcacttgc	attcttcttg	gatgaatata	tctttctctt	50760
tcttctatc	tgggttcgtc	ctctcttttt	atgtgttcag	tttttatgtg	agttactctg	50820
tccatctcca	ttttctcaga	atttcttccc	cctccattga	cgtcatccga	ataactcaac	50880
tggtgcgggt	gatactgatg	acgtcttatt	cttcggacct	tattatcccc	tccttctgaa	50940
tcacagatg	attcagagtc	agatgaatca	gacgaatcgg	atgagtcaga	ggaagaagaa	51000
gaagaatcgg	atgagtcaga	agagtcacat	gagccactcc	caattccact	actaaattca	51060
gcttccaatt	ctctgacgac	atctatatatt	ttagtatctt	tttcttccac	atcttgattc	51120
gtattagggg	atgtcacacc	agtttcggtg	ctttgttctg	cgaggatatt	ttcgaaacct	51180
tcaagaaaag	gaatgtcaat	aaaaccttcc	atttctacga	tctccccttc	ttccaattct	51240
tcctttttgt	tctcttcttg	ttggttctt	gagacgggtg	tttcagtaaa	atttgggtta	51300
tctgtagatc	caaatccacc	agttccacgc	acgttctctg	cttgattttt	gttgtccttt	51360
ttagaactag	agtcaataat	cgttctctct	cccggtgttt	cattaatata	cacaatttgt	51420
tcttctctga	catcacataa	tcttaagaaa	atcaactggg	caatgctggt	tccttctctg	51480
attggcacac	ttttttctgc	actatggttg	cgcagaatca	ctttcaattc	tcctctataa	51540
tccacatcaa	tcgttccagt	aggtacacta	gtgttgttct	tccatgtcat	cccagaacct	51600
gacacaattt	gtccataaca	gccgtcggga	aacttgtcta	ttccatatcc	agtagagatc	51660
ttggccagtc	ccataggttc	gatattccatt	tcttccagag	gaaataggtc	gtaggcgaca	51720
gaaccggcgc	tggcacgtct	gggcggaagt	gcaggttctc	ccccgggagg	ggcggaatct	51780
ataaacacga	cagatgcaga	tgagtccatg	gctggataca	gggagtagac	agtgagagac	51840
ggatgtaata	agtgttcaac	gtctggagga	caagttagta	ccttctctcg	ttcctcccac	51900
ccttttataa	tcgtgaaacc	cccacacccc	cgctaaggca	gtaaaaaaaa	ttgacccacc	51960
atcacccctt	gtccttgaag	gaggaaacta	gaagataact	tttatcatgg	caacattttc	52020
tgaacaggat	cataaaaaat	cgtttttata	tgctaattag	aagctgaggg	aggaagaagt	52080
atacagactt	aaaaatgtct	agccttcagt	ttatgtcttt	attgacataa	aagaaataga	52140
aaatggttgg	gaaaaagaat	tcgggctttt	agtacaacca	ggacagaaat	tagctccttt	52200
cagggatatt	tcttatgact	caagcaaact	tgattgtgac	gcattttctt	gcataccttc	52260
agatatactt	cattctgata	atgaaaaaag	agtaggagag	tgcaactttg	ccgaacacac	52320
ctctgtctca	tttctctgtc	agaaccctga	gggaaaaaca	ttgcgccatt	tcacggcatg	52380
tggtccaggg	tgttaccgga	ggtacaagca	aagagacccc	catactgggt	tgccagtagc	52440
cagaggcgtc	ctgatgcaag	atcacgttga	ccatgaaact	ggaaataaaa	tgtgtgaata	52500
tctgaaccag	agcttagtca	tgtgggctgc	tgtcccttgg	atacgacctg	gagaccttac	52560
tgaaggttac	aacacaacac	acgtgcctgg	ctttgcattc	aaggaggatg	acgaaagggg	52620
ttcaaaaaga	gttaaatatg	aaaaatgtgt	catttcaaa	gcttattgtg	atttctttta	52680
acagtattat	gacgcagact	ctggctcatg	ttatcgatct	ggatggatga	aatttctcca	52740
tttaatgttt	gggcagtatt	ttactaatct	ttcttacaat	ttagctaacc	caaaacctta	52800
caatttaact	ggaaatacat	ggtctgatgt	agtttctgtg	ttgacagata	accctattgt	52860
agatgcgggt	gctgctccgt	ccagatctga	aatggatgaa	attatcacca	agaagaagtt	52920
caacgtattc	ccttctgagc	agacatctgc	tcgccagaaa	gcagagaata	taatacgttc	52980
tcagatccga	gatggtgtag	aaatagacct	atcttctgtg	gatgctttta	tgcagtttgt	53040
taatagggaa	ggtgtagtag	gaacagagaa	aaaatccgac	cgctcatgac	gagtggcaga	53100
cgctgttatg	gatgcggcta	tgccgcttca	agtcattggg	ctggacgaca	gtcaatctag	53160

acgattattg	ttaaaaaata	tgattaaaaat	gagtagaaaac	aacccagaat	atgcaagaca	53220
tttttccagt	tctctcaaat	taattggagt	aactttggcc	ataaaaagat	ctgttttctc	53280
taaaggcgct	tcagctaaaa	gaaaggaaaac	ggccattaat	aatggcgaac	agcatagaaag	53340
aagcagatgg	tcacccgaga	cggtagacaga	agaagatgca	cttttatttg	caagagaaaa	53400
tatcactgaa	gacccaaaac	accctgcgcc	ctttgtggac	attttacct	caccagatat	53460
caattcatct	atcaagagcg	gttcttcttc	ctcaatatgg	aacgatatac	tttcaagaat	53520
ttcttcgacc	agaaaactgg	aagaaaaggc	gagtgttttc	gttaaaaatc	ttgtggtgaa	53580
agtagtgaga	cagttcttgg	acattttaga	aggtaaactg	tttctggacg	gatacgaatg	53640
ggacgataat	attcctctga	tgataggcgt	ggaccaaata	ttgagagaag	tcattaaggc	53700
ggccagtaat	atgtgcgcta	gatttgccct	ctcagctctg	gagtctagtt	tggttactgg	53760
ctttatagac	tcggcgagtg	ctatcacttc	taggttagcc	gtacaactgg	cagccagaac	53820
gttctccgta	ttcttggagg	agtctgttat	agaattttgta	gttgccgcaa	gccttcggct	53880
agcgatacag	gcatttgccg	atctggcaac	tcttgcccg	tcagccttaa	ccgttatttg	53940
aattgttata	ttcgtaatac	aagtacttgg	gttgattttg	gatccttgcc	taggggttagg	54000
ttggtacgat	cacattttta	gcccagagga	tttaaagaag	caggttctag	tgtttaggag	54060
agagtttgca	aaggcaggaa	atgtagatgt	gggtgtagct	caaccagtca	cccctgaaga	54120
aatcgctcgt	atcaacgttt	tccttcaaac	tgaaagaaaac	ggggaagaaa	agaaggaaga	54180
aggagcacga	aaatcaaaga	ttgattttct	tcaaaaatac	ttccattcta	ctcctttgat	54240
gggaaagaaa	agtaagtttg	tctacatata	agaagcagct	caagaatact	tgggaggaag	54300
aacaatgaac	gcttttgggc	agcgtataat	aacagctgct	gatgatagtg	acaccaccac	54360
caccacacaa	gagggaaagga	gggatgacga	aacagtgact	aagaaaatga	ggagtattat	54420
tctagaaaaca	ggtcaaaactc	ttaaggatta	ctcgtctgct	gttaactata	acgcctcccg	54480
tctagattac	gtgggagagg	aatgggtaag	aaatactgcc	ctaaaagaag	agacaaggag	54540
caacactact	agtgataacc	tattcaagaa	aactgtttct	cttgctagta	tggccggcgc	54600
atctctgggt	ctaggaatag	gtgtatttgt	agcgtcccat	attacacttt	tacgtttcac	54660
caatattgggt	ctagcttttg	cgttcgcagg	tctcctagca	tttattgcac	ttatgagtat	54720
atcatatata	aacatgaatg	ctatgggtgt	agtgaattcg	gacgcaatat	acaggtctac	54780
tgctctagtt	ggagatatca	aaacagaccc	cagaagagta	ggaatggtcc	agcgccacgt	54840
aggtgtcggg	gctaaatata	acatgattac	agatttcgtc	tctccaatgt	tagacgagat	54900
cgagagtgac	taaaaagtga	gggttataac	cccccccca	acaaaggctt	taaattagac	54960
tcaaaccaaa	ttcataatca	ttgtgaagac	atcgcaggtc	gaggaaggtc	cggtgtcaaa	55020
aggtcccaaa	aatattcctc	aatcaaaaaa	taccatggat	tcgttgacta	atacagtaac	55080
actcctcgtg	aatgacogtc	ttggaaatca	tcgaacaaac	aaaccaatca	ccgaacaaga	55140
tgtggaaaat	actcttaacc	ttaacagctc	agaaagggca	agtcttttga	agttatatte	55200
tgttttcatc	aaagaaatgc	agtcctattc	tggatgtata	cccaagaaca	agtacacgaa	55260
tgtgcaagaa	atattcgaag	atggactaat	tactttcgaa	tggagagatg	gaacaaaagt	55320
acacagatca	gtttcaccaa	gttccccctat	acctctttct	acaaaaaaat	cgccctcggtc	55380
ctcaccttcc	cctcctccat	cgatgccttc	tcttaaagaa	gaagaatttg	aagaagaatt	55440
tgaagatgat	gaagaaatat	acgaaacaga	tgaaaatgtg	gaagatttca	taaatggtga	55500
tggagaagat	tcagaagagg	aagaagaaga	agatataatt	gttgatgacg	aagaagaaga	55560
gaatgaagag	ggagaaaaca	agtacgttct	cgcattctct	aatcatctga	ggcgccagac	55620
tgccgcgcgt	gccgcgcgcg	ccgcagctgc	tgctgtctgt	gacattgaga	agaaggacaa	55680
aaaccacgca	gttagtgccg	atgactacac	actgtccgcc	gtccagcaac	agcaacaaaa	55740
actgctccag	caacagcaac	agcaacaaca	ccagcagcgc	tcctcatctg	agaaggctcac	55800
ctccacaccc	aacaaattca	acaagttttt	actgccaggt	aatggcttct	cagaacagac	55860
cgagctcttt	gtttgtttcg	atgtggataa	aattgcccga	tataatggac	tcgtggagct	55920
agacatctta	cccattgttg	ctgaatacat	taacacggc	cttggctctga	aatgcagtat	55980
ggaaactccc	ccagtgaaac	cgtgcaggag	aaaggaagtg	aaagatgtgt	ggtgtcagcc	56040
taaaactagc	tttgaaaatg	atgctgtgga	agataaacat	ctcgcattcg	cagaatcgcc	56100
tatacttcaa	aggcctagag	atttccctat	ccctaaaaaa	atcaccgcct	atttttgttt	56160
agacgattct	gtagacatta	aaaacccttg	gggttcgtgt	ccgcttttga	aaagtggatc	56220
aaactttcga	gtgtccgaat	attcgcgtca	ttttaatgaa	ttctcaggag	ttaaaaatga	56280
cgatgacacg	tcttcaaaca	cttgctttat	atactctcaa	aaaaacccca	acattgaaat	56340
tgtatcaaaa	ttaaatattg	aatttgaggt	aatgatggag	ggaattataa	cccatagaaa	56400
agattttgtc	gagacgggca	ttttgagcga	ttcttcatta	gctacggcta	tggcattttg	56460
ccacccaaaa	gctagagttc	gaaatgttgc	attgttttat	ttttctgtat	atttaccatt	56520
ctctaaaaata	actcgcaaa	aaactataaa	atgttcagag	acggataaag	tacatatttg	56580
ttcagacgcg	atcttttctc	ccccaaagtga	taactccta	ataagtgtc	accgaaacaa	56640
taacaacaat	aataacaata	ataccagtgt	gaatactgag	gacagaccta	tccgaaataa	56700
taatataagc	agaaaaatga	ccatcacaaa	ctaccaatgt	atggcatgca	aggaaagatg	56760
cacaaacaat	tgcactaacg	gtaactatcc	cgatcgtggt	aaccagcact	tgctcacatag	56820

WO 01/38351

PCT/US00/28888

17

tgtaaaaggg	gaagatttct	ttaagatttt	aaataatagt	aaagtagatt	cattgaaaaa	56880
attgagcaga	gtactgattc	ccgctcctcc	ctctggaaat	tatacatcta	agttttgtga	56940
tagaagctct	atgtgccata	gcttcttttg	tagagggatc	gaaccagtg	ctacttcttt	57000
ctcatcggac	agttttgaaa	agactaaact	tgttttgtat	ggtaaagtcg	ttgacgttat	57060
caatagttaa	tctgccataa	aaacttccca	taataataga	atcagggctc	tttttaactc	57120
tgaggaaaaa	gataataaga	ctatccccct	tagagccgaa	agtgcacaaa	atgcattcaa	57180
ggatatactt	gttcacgaat	gtaataaaga	acgagctggt	tcataatttg	agcaaaaaca	57240
attatcctct	aaagatgggc	atctatctaa	caagtgggtg	atcgaactta	atgacttgaa	57300
cattatgttt	gagaaacacg	tggaagattt	ttacaagaaa	tggttctaaag	taaattgatgc	57360
agaaatctta	aaggatattt	ttaatgattt	tgaaaaaact	tgtgataaat	acaaaactgc	57420
caagagggca	attattggag	cacaagaccc	ttctacttct	actccctcta	aaaaggagaa	57480
tggtatcact	aggattatta	gtacattatc	cgaatttcat	tcaaaaagatg	aagctacagt	57540
aaagtccctt	ctcgacaaaa	caatgctctt	gggatcggag	acaataatgt	ctggtgttag	57600
atgtgttata	cgtaacaata	gtgtgttttc	gggctttgaa	aataagaaca	ctaataataa	57660
ttgggaactt	gagattagac	actatgtcat	ctctatggga	ggtgctgcag	tgacaaagat	57720
ttccgatgaa	gattttggaac	aattcacgcc	tgtaagaggt	gctgtctctg	tcactacagc	57780
acctaataat	aagctacctg	taggggcaca	tcagacatgg	aaggatgaac	aaacactaaa	57840
aacaaacact	aaacgtaata	gtctatatga	ctcttacaat	tcaaaaagga	ataataggga	57900
taataataaaa	ataaaaaatc	gttcattaaa	actatcagat	tttaattgga	gaacacccaa	57960
tatctctatt	caagaattta	atgcaataaa	agatgatgtt	aacaagaaga	ggtacgcaga	58020
agtcgtggcg	tcagctgctc	caaagtcacc	ttcaccaacg	agcagcagca	gcagcaacag	58080
caacagcagc	agccctctct	tttcaccgct	ctcaccaaca	gtgaagaata	gtaataataa	58140
accattgtat	attcctcccc	ataaaaagaat	gaccactact	gctgtttgac	caatttatgg	58200
gttattttat	aatgttgtgt	gatattttat	acaataaaaa	tatataagaa	aacaacccaa	58260
tgatttttga	tatccttggt	ctgtgttggg	cggaaattcc	tccgcgcccc	ttatgatgaa	58320
aaaattgcta	gaggccttga	atcctgggtc	gcacaaggtt	aacaacaaaa	aaaattacct	58380
gtatgtgcaa	gtcatatcct	ccagtcccac	gcattgagaac	ctgttattgc	accacgattc	58440
gtgtaacaaa	gtcgtatctg	cacatgtgca	taaattcctt	cacgctaaaa	tgaacgacaa	58500
aattgcccgg	agcaatgggc	gtgagagccg	gttcctggaa	aacccacact	ccgttccacg	58560
cactacctca	ctattgcacc	attttttggg	gtataaaaaa	gcacggagag	gtagctaaaa	58620
catccactcg	cagagacatc	accgtcaaga	tgacgtctcc	tactagtgtc	ctccgcgttg	58680
ctactactgc	cctgctgttt	gctgtgtgtg	ttacggctac	tgtgatttct	gaagctgggg	58740
gaactaagat	agacgaccgc	tgccctctcc	cgccatgtga	ctacttcccc	gaacctgaat	58800
gacgtggaag	aagatgtgtt	agagctgctc	gaagggggaa	caacaacaat	aacaacggaa	58860
cccgtgatgg	agccaaagtc	tttggcgaaa	tcgcaaatgg	actgattggg	cgtaagaggc	58920
gtgaagcagt	atctctaaca	ccagtccatg	aagatatgcc	agatttcttc	ccctaccctc	58980
accctgagca	tcccattgga	gggcgtctac	cacgtgaagc	agctcctcca	cttaagggtg	59040
cacttggacg	taaaaggcgt	gaagcagctc	ctccacttaa	aggtgcgctt	ggacgtaaga	59100
ggcgcgaagc	agaatccttg	gaggaagaac	ttgtgtctgc	tgaagaagaa	cgtgaaaagc	59160
gcgaagcagc	tccccacttt	aaaggtgcac	ttggacgtga	aaagcgcgaa	gcagctcccc	59220
cacttaaaag	tgcaacttga	cgtaagaggc	gcgaagcagc	tccccacttt	aaaggtgcgc	59280
ttggacgtga	aaagcgcgaa	gcagctcccc	cacttaaaag	tgcaacttga	cgtaagaggc	59340
gcgaagcagc	tccccacttt	aaaggtgcgc	ttggacgtaa	gaggcgcgaa	gcagaatcct	59400
tggaaggaaga	acttgtgtct	gctgaagaag	aacgtgaaaa	gcgcgaagca	gctcccccac	59460
ttaaagggtgc	acttggacgt	aaaaggcgtg	aagcagctcc	tccacttaaa	ggtgcgcttg	59520
gacgtaagag	gcgcgaagca	gctcccccac	ttaaagggtgc	acttggacgt	aagaggcgcg	59580
aagcagaatc	cttggaggaa	gaacttgtgt	ctgttgaaga	agaacgtgaa	aagcgcgaag	59640
cagctcccc	acttaagggt	gctcttggac	gtaagaggcg	cgaagcagct	ccccacttta	59700
aaggtgctct	tggaagtaag	aggcgcgaag	cagctcccc	acttaagggt	gctcttggac	59760
gtaagaggcg	cgaagcagaa	tccttggagg	aagaacttgt	gtctgctgaa	gaagaacgtg	59820
aaaagcgcga	agcagctccc	ccacttaaaag	gtgctcttgg	acgtaagagg	cgcaagcag	59880
ctccccact	taaaagggtga	cttggacgtga	agagggcgca	agcagcagca	gcagctatgc	59940
ctccccctga	agacgatctc	gacttctttt	acgcacctgt	tgctttgcct	ctacatggag	60000
tatggaaaagc	accagaacct	acagggttaa	atgtggctga	acatcacata	tacaagttaa	60060
ttggataatt	gaacaatttt	gacttttttag	tacttaattg	ttattaattg	aaacaatttg	60120
actctttcct	acctaagggt	tattaatttg	tacattgtta	acacttgctc	tttgctcata	60180
gcagacatac	ctgctatttt	taacgctttg	tttaacgtgc	tttttacaat	tgatattggt	60240
acatgtttgt	gacctcttga	ataaattttg	ttagcatcaa	tattttattt	atttttatcc	60300
ttttcttctg	tggttagaatt	agaaggagat	gtggagggtg	ttgatgtagt	agatgctgtg	60360
cagagttgct	cctggtgttg	ctcctcttct	tgattcacat	tgctctcttc	aacatactcc	60420
tctaataccaa	agcgtctctg	aattttctct	tttacaccct	caccacaatc	tcctatcgaa	60480

gactgtgcta	atataat	ttt	tgtagtagaa	gtagagaata	atggtacaac	gttattgatg	60540
ttgttacttc	tgcccatagc	tggttagtatt	tttgaaggca	tgcccttaaaa	aagggatata		60600
tctttgtcct	tttctataca	aaaggataat	cgttcatccc	catgaaaaag	tgtagatg		60660
ggcaatgtac	acctatagaa	cagtgttatt	tcttctatat	cttccagtat	ttataccaga		60720
ccctcttgcc	ccatagtttc	gtatctccat	tttactattc	ctgtctctct	ggggaaccca		60780
ttctgagcca	gggtgtctgc	ccataatttc	tgtgttttta	tcaacaacaa	tatcagctga		60840
aggggcaaaa	ggccccccat	cttccagtgc	actattccca	ttagaaggaa	gtggactagc		60900
tgccataggt	gacaataatg	aagggagtgg	tgaacacaca	ggataatcat	ccacaccctc		60960
tgtcactaga	gattctagat	ccgggggagg	agggtgaggc	aacataattt	ccatctgcat		61020
ttcttcccc	gacgtggacg	gttgctgttg	ttctactcct	tctacctcta	cttctactct		61080
tcctccttct	ttattatttt	ctgcttctag	ttccttcatt	accatttcca	cgcctgttgc		61140
tagtctcatt	ttcttattgt	tattttttat	gcgttcttta	gcttgctttg	ttacttcata		61200
tgtaatatca	gatgagtagt	atttgcctgg	acaaataagg	ttatcactca	tggtgtgata		61260
attagtatca	gaattgccat	gtgcattatc	ttttgttata	aactccttgg	gtgggtttatt		61320
tgaccagccc	gtgaaagtat	ttgccagaag	agccgagggt	gatgaagagt	tttgcttaac		61380
atacttttca	tcgatagttt	tattaaataa	tgtaaactaaa	ctggccgaac	acttactatc		61440
ccggcaattc	gcacatgcat	tattattctt	ctttttatta	tccattgagg	cctcctttgc		61500
agtttcaatt	aaagctgata	tttggcccat	gattttaagg	ctcgcacaca	gtgccatttc		61560
ctgtcctcct	ttctctagtt	gaattgatcc	ccattcatta	aatttttctt	gtagagtcag		61620
gagactactc	ttgagaaaag	agtgtaaaag	tgtttgagta	gttatacaat	ttttcaacgt		61680
ttcagtgtat	gggtcgaaac	cgaatttgat	ctgatttctt	acactatcca	gtaaatcctt		61740
gtttaattct	tcctgtctag	aggaagggaa	agaaattata	tcattagctg	gaattgattc		61800
agaaaagtga	atatttttaa	gtacatttcc	gtctgttgaa	gaagggtgct	ttagatcaca		61860
cgtgctaagc	attacatcta	atgataagaa	aggggttaact	ctttctcctg	tatcgagatt		61920
tcctacaaca	ttttctgggtg	tcaagagatc	gtctgtttct	gtcactccca	ccaatctctt		61980
gtaatcttcc	tcagcaacag	tttggttggt	gttatagaaa	gaccataaag	tctgaaaaag		62040
tctcctcgca	gtctcattat	ctgcatcatt	agctcctcct	cctcctcctc	ctcctgccac		62100
atttaaaaga	tctaggtagc	tactcatctg	ggagggaagg	ggagggttaa	gggttgata		62160
taaagaggct	tgtctgaaga	cgcgagtgat	cacaacttca	acgatggacg	tcgagttcgg		62220
tttcttccac	ggctgtctct	ccaaggccct	tctccagat	gaaaaacatc	aaccgcgttat		62280
aaggcgtctt	tgtgcggatg	attctagaaa	taagggagag	gatggctgct	gctcgttctg		62340
tggaagaaga	ggaacaggag	agagcaatc	tgcgtgcctt	gaacaactaa	tagacgtttg		62400
ttcctttata	ggaactgtct	catctatttg	tacaattatc	aatttctaac	tttcaaccag		62460
ttgttctaga	ctacaaaaaa	cgtcagacag	ttatgcggca	ttatcccat	ctagttttct		62520
ggatgtggta	tatccaagtt	tgaagaaaa	aactgaagac	gtattgcctc	attctttacg		62580
tgccatttgg	aataaacaac	ttccaaagtt	gtatgaaaa	actcttcaac	ccatagaaga		62640
agaggtat	ggttataagg	attatgttgg	ttcaattgaa	gacgacgaca	atgttgatga		62700
tggtgacca	caagaacaaa	tgtattattga	tgaagaatct	tataaaacta	ttggagaaaa		62760
atcaaccatt	gaactgatag	gcatgtataa	caataacaag	tttggtaatg	aatttataag		62820
gattccttta	agagaaactg	cggtgcacgc	acaatctctg	aggtacgaca	ctgaagctaa		62880
atttgttaac	cacaaggact	ctatacctct	attttatgaa	aacagcacgt	gcacatgtaa		62940
ggaacgtctt	attgattttt	ctgagagaca	actacaacaa	ctaaaacaag	atggaatgga		63000
taaaccaacg	gacaagtaga	gaggatcttt	tcaaacagta	tacagggaat	gcagtaatac		63060
gttcggcagc	taagcaagca	ctggctattg	aaaaacacgc	agcagaaaga	agaggagaaa		63120
aggcatggac	gacttcagca	gcagcagcag	cttcttctaa	ttttaataat	gtacaacaag		63180
attatactga	tgatgatatt	acacaagtgt	ctattgcaaa	cagtgttttg	aataaccctt		63240
ttttaagag	atatgcaaaa	cttatagata	atttagcaat	atcttcttta	cctcctgata		63300
tagaggatga	tgtcattata	cacactagag	atgcctccaa	ctctacagtc	agagtagatg		63360
gagccaatat	ctatttctgcc	ataattgacg	gtgatttatg	tgtataccct	aaacaatata		63420
tatctgataa	agtgtctgtg	ggttctctca	accgggaaaa	ggcactgttc	tataatagct		63480
ccaagaataa	gtggacgtat	ggatgttaac	taaactttga	tatcgttgac	gctgccatca		63540
tgaaacaccc	cgactacaag	gaagagacta	catctacaaa	acatatagct	aaaatatttg		63600
gtatcggagc	atcggaaaaa	ctgaacatta	cccactattt	aaactacttt	atccaataaa		63660
tcaccatgtc	tatatacagc	ccagaaatgt	cagccatgat	gaatgagaat	gaagaaaaaa		63720
acgttctcag	taataggagg	gctgtaatat	tgatcttttt	cataggtgtt	gtcattttat		63780
cactcttggt	cttaatcaga	ctagatatag	atatcaaaata	tgggaatgtg	ggcgagagag		63840
ccactattgt	aaagaagact	agtcccattt	ttaaggaaacc	tatccagaca	gccatcaact		63900
ttaaacaaca	aaagatttcc	cacattgaag	atggaggaga	agaaaagtgt	gttatagctt		63960
caaggactgt	tagttctctc	cagagagatc	ctcgtgtact	tttaaccagt	tcatttggat		64020
tgtataatag	gacagattat	atgaaaaaat	aaaagtacca	aataccatat	caattgtttt		64080
atttcatcat	tagttaacat	ctttttaaat	ccacaatcag	agctaaaaca	tagtgcagtt		64140

atacaataaca	ataatacatt	tctccctcca	atgatacaat	gttgatgttc	tcctcccata	64200
tacttttcaa	tcaaagccga	gtcttggtcc	ttggataatt	tatccatgac	cagcttagaa	64260
tttctccagt	acatgggacg	aaacactaac	ctcatactag	gattcgtatc	attgtgggaa	64320
gggtttatac	aacttcttga	tgtctttaca	tcctcagtag	gcatttcacc	caataaatta	64380
aacacatctc	ttataataga	aacagactca	gaatcatact	ttaaatcttc	tatgcattca	64440
cctcctatct	tagattttatt	aaagggtgca	aagaaagggg	tttcttcctt	cttcttcttt	64500
tggacatcat	ttttacctaa	tacacttaaa	cctatctggt	ctcttacaaa	ggactgaaaa	64560
ttgatatacta	tataatgtac	atcttcaaat	ttttggatat	tgcgattaac	atttttgtat	64620
gattctggaa	cagggatctt	ctttgcagtc	gtcatcattc	tggaaacacgt	gcaggcaggt	64680
tctgggacaa	cacccatgga	tagggcgata	gcattggcca	gtttcactag	ccacattttt	64740
gtttgtggat	taaaagtctt	cttctttcca	aaacgaagaa	agaagacgct	acatacgtct	64800
ccttgctcgt	caccaattac	acgtattgtg	tcccttttgt	ttaatatatt	gtaggcaaat	64860
tccagtagta	cttcttttcc	ttcaacaaga	ggtattccaa	gttcttcaaa	atccttttct	64920
aaataatcca	cactctcctc	taaagacaag	gaagaggatg	tgcatttggg	cagcaagatt	64980
tctgtggtgc	acatattatt	attcaaacct	tgagaagagg	acatttttct	ctttactggt	65040
caggacgagg	agaattggga	ggagttatgt	tagtttagtg	tacgtcttgc	cgctgctgat	65100
gctggaggta	gaatgaagac	cacagcgtga	ggggcgacca	tggcttagat	atggccacgc	65160
cgagcactgg	catgaattgc	ccctgagagg	gtagccattt	agtgttttca	ctcggtattt	65220
ttccctcttt	attgagagtt	gccaaaagtt	tattatagtt	gtctgtgcat	gtcgatacct	65280
tgtctttaa	atcgtccaat	acttttgga	tcctctccca	gttcttgctg	tttatttgct	65340
tgattctggg	agtccttttt	agttcaggca	tatactgcaa	cattgtgtat	tctagtgcag	65400
cagacataaa	atttagttcta	tctgtgcata	ttttatcctc	ttctgtttta	ccacaaaatt	65460
cactatgtat	tatgggcatt	aattgctggt	acacaatagt	attatcctct	agaccgcaca	65520
tgttcatata	cacgataact	aaaaggttca	tgggtttttac	actataagga	acaccagaga	65580
atcttttcac	gtcgatgcct	tgaactgggt	catagagcac	cacgttacta	ttgttcccgt	65640
cttctctctt	attggcgctc	attaaaactc	cacttagacg	tgataataaa	ctattgtcat	65700
tgtacgttat	attgtcaaga	gaaatgcatt	tgtogaattc	gttaacaatt	tctgccactc	65760
tttcttcac	catcacgggt	tgtttattgt	taagttttga	tgatacagaa	catcttgaat	65820
gttcatcctc	ttctatatta	tttttagcga	gcgatgacgt	catttcgttc	gtggctgcag	65880
ttacagcttc	taaaaagtca	gaatctttat	actcgaacgc	tttcttgcc	tcatctaacc	65940
agtcgtcagt	ttctagatct	gcctcctctt	cctctttttc	accttctctc	attttttctt	66000
cctcctctc	tacatcatcc	atttcattga	attttttaaa	gaaatctaaa	aagtcactcg	66060
tgtctgttat	ttgttcagtt	tttaagattg	acgtctgctc	agtagataga	cctgggtattg	66120
acgatactat	tcctcctcct	cctccatttt	cttctcccaa	tgcttctgga	ttcgtatcta	66180
aattattacc	aaaaacatca	atcctttctt	tagtcaatat	tcccctatca	ttaattatac	66240
caccttcagt	gtggctctca	aaaaatccac	gaattgatcc	gaaaaactgac	gatgataatt	66300
ttttagtaac	aggtgcactt	tttctgggtg	cattaaacct	acttttctta	cgggtagatt	66360
tttgagggtc	attaccactt	ttacgtttat	tattagaata	tgggataaga	tcttcagtat	66420
cagattctac	atccgaccct	gtattttttac	tctcattatt	attagcaccg	gaaagatgta	66480
tactgatata	tgataagggt	gttattgtac	caaaaacacg	ttgattaaag	gaccccatg	66540
tcaaataatga	catgacattt	gocgctgcat	ttacagccag	ggctctatga	agtgcgtggt	66600
tcaaactaac	gtcagtttcc	aaatcaatca	agtgtgttgg	aggaaaacac	aactgaagcg	66660
ggtctcgtag	aaaatctata	gcttgagaca	gatagaaaaa	tttgattgtc	cagaaaaggaa	66720
gtactagttt	cttatcacat	tttctacaat	tactttttaga	caaattcccc	atagaaatat	66780
taccaacact	atcatcacag	tctttattat	tcttcttagc	ttcttggtca	ctcaaaaaaa	66840
tctcccatcc	gccacacaaa	gtacagtacg	tgttttcaaa	tactgcaact	tctttcctta	66900
cccatgtgtc	gtactcacag	ttactaacia	ttgcactatt	gtagtctgat	ccttttagtta	66960
gcataatcaa	tagtgtactt	tcaggtgaaa	gcattgagcg	cgattctgcc	aaattatata	67020
tcgtatttgg	ctttgttctt	cttttagctg	caaaaccacc	caactcgttt	gcttgatttt	67080
caccggaaga	attactgttc	atgcacctaa	agaacttgac	atggccaatc	tttaagaacg	67140
tttcagagtc	cttaaattgt	gtaccagtaa	gacgacacac	gtttttatga	tgtccaacca	67200
tcaggttcca	tttcaccagc	acatcagaat	cgtggagaaa	aagattcata	gaaagtgttt	67260
gattttcatc	gtttgtgtgt	tgtctcaaaa	cattcaataa	atcttcccac	ctcttggtat	67320
cttttctgtt	accagacata	ttataagaat	cgtactccaa	tatcttctct	ttttctagta	67380
gggagaagtg	tttcccagga	gtctccagat	tcaaaaggca	tgatgttaaa	cgtacgatata	67440
catcctctgc	ttcgggtgaat	gatgcattag	aaagggtacat	gcatttgctc	cccgtagtct	67500
ttagaacagg	gatcctcatt	atgtccagta	atgcagatgc	catcagacct	tgagggaagac	67560
gcatacatcat	ccgctgcac	gagtagaata	gtagatcggt	ggtcttcccta	ttgagggaat	67620
tttcttctct	tgattttgcc	ttttcctttac	ttctttcttg	tctgatctgt	cgtttggtct	67680
tcgtcatcat	cttgatggct	ggagcctttc	catccgtaaa	gagtacagta	tacctccttt	67740
taagggtttt	agacttgagg	agattttttca	tcactcggat	catcctaaaa	gcgtagtaga	67800

agatgacccc	ctccatctgg	ctagcaaaaa	atgcatcttt	gaatctggat	ttcaaaactg	67860
ctgtcgggga	agaattgaat	gaagtagtag	gcacgttcc	gaacaagaaa	ctgttagtca	67920
aaaacttcct	cttcttcata	acatctgtgt	gtttgttgat	ccagtaatcg	agaccctcat	67980
agtcgtccat	tttgtgatta	gaaactggag	gagccatcat	catgtcttcg	attggagggt	68040
cagttagcca	atcaacttcc	tcttgggatg	aagggattat	aataggcact	ctttcaattc	68100
tagccttaat	tttctcgtgt	ttatattttc	tagaaaaaga	ttcagaaata	accttctttt	68160
catccctgca	gtaaaaattt	tcataaatat	gtttcttggt	ctgcatccgt	gaatgcatag	68220
tatcttcttc	cctgataaca	ttaaaaccgt	catgtaaaac	gtctattatc	tgttgatttg	68280
tagtattgat	aaacaggatc	atctgaccca	tatttttctg	tcccttcaac	atgggtggtg	68340
caatttgcct	gccaacatg	accagatctg	atcacattata	agtgggcaaa	tcaaccatgt	68400
cgagtgtgta	gtcacaggta	ggtgaagaag	aggaagattt	cttgccctct	ccaaaatcca	68460
ggaaagtctt	ggagttgatg	ttggcaagag	aagtggcctt	gagaagttag	ctagttgtgc	68520
aatcttcttc	tcctgaacaa	taatcctcgt	cgtcttcgtc	atcgtaacaa	tcttcataatc	68580
cagttttctc	cttgatgact	ctcccattct	gtgcaccgta	cgcattttctg	tacacttttt	68640
tggcttctgg	gacacccatt	tctgttgatg	ttactccac	gggatgttag	acgataatga	68700
atgtcactg	atgggcgcg	taatttttac	acactctgcg	tagtctggtg	tttttaccce	68760
tcctagaaga	acttcgtgga	gctgctttaa	aggttgtagt	gcttggaaatc	tgcagcaaac	68820
gcagctctct	acagagactt	ctcggagact	tgccagaagc	gagtttctctc	ttcttaatag	68880
gggtactttc	ttcctcttcc	tgttggtgtt	tgtagagtgt	tatagattca	gtagggaagt	68940
agtttctaga	tcgaagaaca	gtgtagtctt	cttctctttt	cacagtctca	atagaattgg	69000
acacctcgtt	gaacacttct	agggccaatt	cccacacacc	ttgttcctta	agtgtctcga	69060
ccacgtaatc	aaattcggat	tcactcaata	actgtaagga	ggaagaatct	ccatccatga	69120
tacgaatctg	gaacacctgt	ctcctccttg	actgttaact	gatgtctgac	cccttgtaac	69180
cactttactg	atgttttata	ccctctccc	acggaagatg	tatacgtcac	aatttttttt	69240
aaaaaataac	gttatacacc	aaaaatgagc	gatacagggc	agatggaaga	aaataggcct	69300
gtacaccaga	aaaggagacc	tggagatgaa	gaagaggaag	aaactggtag	tagtaatgtt	69360
ccatattatg	ccaacttttg	cgatgacgcc	actgtactcca	tgtacactgg	agaaggaaaa	69420
aggggtaaat	ttgtattaga	gccaccta	gaaagaagt	tacaaagggt	gcaaaaacca	69480
cctaaagaaa	aggaggaaag	ggaacaacgt	tctaagtctc	ggacacggag	acctgggtcaa	69540
gagtttgaa	agaaagtgt	acaagatcga	tcgcgagaac	ggtcagaaaa	acttgggcaa	69600
aattttggcag	agaaaggatt	gcaagaacgg	caaaagaaat	atactccaaa	ggtagcacia	69660
acaatgacaa	aaaaaataat	caggtttctg	gaaggaggaa	gaaaattcaa	ggcgccgcaa	69720
cagcagacat	ctgacaaagg	tgcagcaacc	aatgttcttg	aaagggaaga	aattgagatg	69780
gctgcagaaa	gagaacaacc	agtagaaatt	acaggagata	ctatatagg	tgggctagga	69840
gaagaagatg	acgaagatat	gggagaggat	gaattaaacta	tacaacattc	atctatggct	69900
gtatcacaac	ccgttcaaca	aatcgttgtc	agttctccta	taccgccaaa	gcccactagg	69960
cccgctccctg	atatttccat	acaagaagat	atagtgggga	aaaatattag	ccagttacca	70020
ccattaccac	ttgatgatta	tgaggacgaa	agaagcgaac	atltgtacga	agaagtgaat	70080
gattttcttag	tggcaccacc	aacagcagca	gcagcagctt	ccacaagacc	tcccaggcct	70140
aatattcctc	ctccacctcc	tctgttgtt	gctgttgacg	acgaaacctt	gaagaacttg	70200
gcttcaattg	cagccttgga	aaaggaggcc	gaggaacaaa	gagcgccgcg	agttgaaagg	70260
gaaagagaag	tagaggaaca	aagagcggcc	gctgtgtctg	ctgctgtctg	tgtgtccgcg	70320
caacgggaag	cagacgaaaa	aagggaaaga	gaagcagagg	aacaaagagc	ggccgctgct	70380
gctgtgtctg	ccgctgtctg	ccaacgggaa	gcagacgaaa	aaaaggaaag	agaagtagag	70440
gaacaaagag	cgccgctgtc	tgtgtgtgtc	gccgctgtctg	cccaacggga	agcagacgaa	70500
aaaagggaag	gagaagtaga	ggaacaaaga	gcggccgctg	ctgctgaaag	ggaaatactt	70560
gccaacaac	ttcaagaaat	gaaagaacaa	atgcgcataa	aggaagagga	gaggcggaaa	70620
gaactagcag	ataaggagga	agaaaaacgt	cgagaactag	cagccaagga	ggaagaaaaag	70680
cgtcaagaaa	tattagctaa	agaagagcaa	cttgaaaaaat	tgaatttcca	gttgggtaca	70740
gaaatcacgt	ccaaaagagc	actcgaacaa	atgttagaag	aagagaaggc	ctcacgctca	70800
cggtccgcag	ccagtgcaca	gttagcgatc	caagcaatag	aatatgagga	tgaacttcct	70860
caggcagtcg	aacctcaagg	acagttagtt	cctatggata	cggattttgt	cggaaaaatg	70920
tacgatctca	ataagaattt	agaagtacag	aataatacat	taacttctgc	atltgaagac	70980
gtgaacaaaa	caaacgaaca	gaaccaattg	ttgtctcaat	cccttgaaaa	atccgctaaa	71040
gccattgaaa	aattaactag	tcaaaaacat	cttctgtgtg	atgatcctgc	ttttatgcag	71100
agaataataa	cagagaggga	tttttcttta	aagaatctgg	gaaatgttta	caaaagagtt	71160
ctcggggctt	atltttacatt	aaaaagggac	ctlttttaaat	cgaaggcatt	aattacagat	71220
aaagaatcaa	gggatctgga	ggtgcgtcta	acagatgtat	cgacagatct	cagggctaatt	71280
gatctcaata	caatactgga	aaggttggtg	gtatccgtta	acatacgctc	tgggtgaaaca	71340
ttatacacta	aattttacaga	ggcagacacg	gcattagcag	atcaagttcc	ttcgaggatt	71400
gaaataagta	acagatcaag	atctgcctta	ttgccatttt	catctgcagg	tttgataact	71460



aatttttacta	atagttccga	caagtacaat	gaaatagtga	accaactaag	cagtataaat	71520
gaggctatga	atatttttgaa	agaaaatatt	gtcccaacat	tgaaccaa	caaaattgat	71580
gtcaccaatc	tattaacagt	ttcaagctct	cgtcaatatg	ctattgaaga	aagggtgtat	71640
tctgatgtgt	cccgaatgga	ttctgaaata	agaaaattcc	tcgctataat	gaacagtaaa	71700
atatccccctt	attttaaagg	cgattggacg	gatgaaagac	aacgctctat	tgctgacagt	71760
atttcctctc	agataaaaac	aaacgataaa	attaaagaga	gtgttgctac	actacacgat	71820
atcaatacaa	cctcaagaat	acgtagtaat	cccctcctgc	acaaatcctc	agttttatca	71880
tctccagact	ttttaaatgc	tgtaaacgac	tttagaaatt	ttctcgatat	ccaaggaggt	71940
tctcaattta	cttatgatgt	cctttcaggc	caaaatattg	atgacctttc	actggcatca	72000
aaaaccactg	aaaaggttac	agaattgtgc	ctcgaattat	ccataatttt	agacgtgatc	72060
cataaaaatg	ctttgagttt	aaatttgctt	gcaattactt	accccgagg	ggaaacatct	72120
atggaagaaa	gtggttcatt	ggctgttgac	attagacaag	aaattggtaa	gaacatatca	72180
gattctagcg	ctgaacttag	tcgcacactg	tcagaggcgt	tacaaatttt	tcagcaacaa	72240
cagcaacaac	aacagcaaca	attccaacag	cagctgttac	aacaacaaca	ggaccaacaa	72300
aatcaacaac	aattattaca	acaacaaata	gaagaacaac	aacgggttca	ggaacagcag	72360
caacagcagc	aaagggacca	acaacaacag	gaacaacagc	aaagggaaac	acagcagcaa	72420
cagcagcaaa	gggaacaaca	acaacaaagg	gaacaacaac	agcagcaaca	gcagcaacag	72480
agtgaccagt	ttcgacaaca	attattgcaa	caacagcaac	aatttcagca	attactacaa	72540
caacaaggaa	gaagaagagg	gggtgacgat	ggtgatgaag	aaagagagga	aagagaagaa	72600
ggggctgaaa	aggatgattg	tgtgcgtaag	ggtgcagaat	cagtagcgac	aaaatatact	72660
gctgacttga	ctaccttatt	ccaacgagaa	gaaaaact	tccaatctaa	aatagcatca	72720
gcaaaaattgg	gaacccttgt	ctttgccacc	cctccttcac	ctatcatgaa	cttgacaaaa	72780
ttgagagagg	aatattccac	attcacaacc	cagtgttttt	caaaactaac	agctgaaaat	72840
aatagtatta	tgcgattttt	ccccgagagg	attgtagaag	tatgcaagag	taagaatctc	72900
aatttaaatgg	gaaaatactt	gtacattata	actaccgcac	aaacagaaat	ggaagatcga	72960
gtgaagaata	tattgtctgg	tattttcaat	caaattgaag	agttttcaaa	caatgtaaaa	73020
caacaacaac	aacaacaagc	tgcttcagct	tcttctacta	atcctcctcc	tccttctact	73080
ccttctacta	ctcctcctgt	tacaagcatg	caagtttgtg	agttggatga	tcaacgtacc	73140
ctagaaaagg	ctgccatagt	agaggcaatt	actctggcca	atgctgtact	tcaaactaca	73200
aaatccgctt	cagctccttc	cacggcggct	gagcgagaaa	ttgctctaaa	gctagagaa	73260
gggaaaacat	ctatccgtat	ggaaaaagtg	gatctaagtt	caggagctac	tggtgtttcc	73320
gaccaacaaa	aattggatcga	cgaaagtact	tccaaacaag	aattggaaga	tttcattgca	73380
gaagaaaact	ttgtagaaac	tgacataaat	ttggatgaat	ttggattaat	tttggtgcc	73440
aagaagaacg	atccgaccgg	tgatgccaat	cttaggctcg	tgaaacctca	tggaataaat	73500
gtgcagtctt	tcccatatta	cgtactccgc	acatggctag	gagaaacaga	tatattggat	73560
gaagatactg	tacatcctga	atatttccgc	caatacattg	atcgcaattg	gaagggtggaa	73620
gaacatgagc	gtgaagatac	attaaaggca	ctgggtgttt	ctttatcaga	tacgttggca	73680
cacatcaagg	actactattc	tcccagtttc	aaaaatgatg	catcaaaatc	agtaccattt	73740
gcgttgaaca	ctctatttga	caacatattt	gctatcgacg	gaggaatgat	ttctagcctt	73800
tcaagaacag	cctttattta	ccgtaaattt	ttaaggcaat	ctatgacaga	taaagaggtt	73860
gctcaaggac	ctgttcggtc	tcaactgtgt	gaagcgacaa	tagcgtctct	tttcacggca	73920
tgtagtaaac	ttcttcgatc	gtctccctta	gccgataaag	tagaaccacg	ccttcaagaa	73980
aaattagcgg	ccgctgctgc	tgtagacacg	tcaaccggag	acatgtttcg	tatacgagtt	74040
tgtcatctca	tgtacaattt	tatagtggct	tatgtgaacc	tatgcaataa	tcgtataaac	74100
tatacgttaa	atgtattgag	agcgtcgggt	ctggcaacaa	aaaaagtcgt	ggccggtaaa	74160
acaactaaag	ggcatacatc	ttcttcccac	cggtttggtt	cttatgatgt	cacatatgat	74220
ttttcagtat	tgtacaagat	tcttcaacta	caaaaacaga	atatctcttt	acttttagag	74280
aagggtttca	atgcatggga	atcggtgtgt	gccgccatgg	cagccttcac	tgccgacctt	74340
tctctctcca	tctcagatgc	cgaccaatcc	atactcttcc	cactagaagg	gggagagatt	74400
gttattgaaa	aacatgaaaa	cgatgcagaa	aaaaatgtcg	atatggttca	agaattgtgg	74460
aaggaaactg	cactcacact	catggcaaa	gaactaaatt	catactacaa	ctggttcata	74520
tacagtaagg	ataccgatat	ggaaaagtgt	gcgagagtgt	gcaggatgat	tatcgggtata	74580
gttaaggctg	ttctcagatt	aacaaataaa	gccgagagtt	tagtagacac	taatgctctt	74640
tctgatattt	tcaagctacc	tgtgattcca	attgatgata	ccaaaactct	ggcaataaat	74700
attgtggtgt	tcactttgaa	taatgttaatt	aaaccatgga	tggttttcgtt	caagcaaatg	74760
tttagacaga	aagacggggg	agttttcttca	gcctatattct	cgtttcaaaa	tatccaacaa	74820
caaaaacatc	aacaaactgc	ttccattttta	gatgcatggg	cttgccgcgc	cggtaagtta	74880
acaaaggcag	cacacgtgtt	tatttcagga	tatgaaaatc	atatcaaatt	aaaaaaaagt	74940
gatcttttat	ggggtgcttc	aatgaaattc	cctgccgacg	gacgagccac	tgtggtcgaa	75000
ggatgggccc	aacaatacaa	taacgaaagc	gtattggaag	attttaccga	cttttcaatt	75060
gaagttaaag	cacctgcctc	tggactctta	ataccgccag	atcctttact	ttcttccatg	75120

ttcggtaaaag	gaaatggtgg	aagcagcagc	agcagcagca	aggataatac	aattatttga	75180
aagggaggat	taatttttaa	ccgccaagtt	gttggaacaag	aacaagcacc	accaataaac	75240
acgtcgtcgg	atactaaaaa	aataagacgt	gatgcaaata	ttgaaccaat	aatagggacg	75300
ccttatagtg	taattaaggc	aagtaaagga	gtatcaattt	cagtactgga	tgattttcaat	75360
gaggatagtc	cagaagattt	cgcccttaaa	acttccatca	tcaatgatgc	catcacgaga	75420
atagggcaac	gcatgactta	tacaagacct	atatttggatc	atcaaacaca	gaaaaatata	75480
cactacagtt	cacctaaaat	tatttctgaa	ggatcagatt	taaagaatgg	acaacgttca	75540
ggacaatctt	gggtcccttc	ttcatcatcc	ttgactctgg	cctccgattg	gaatctacct	75600
tctctggagc	ttttatatag	agaacttgcc	acaaaacaag	tagagaagga	agaagaagaa	75660
aagagcgaaa	gggaagaaga	taaaggacaa	aaacttaatg	aaaaattatc	atttgtcgtg	75720
aataaagcta	tcggaactat	ccaacaacaa	catcaatatt	ctgaaagggg	aggagggaatg	75780
aagaggatc	agcaacactc	tgtgatcaa	gctagtaatg	gtggcataga	tgatatagaa	75840
cttatgaata	gtaaagatgc	tacttccatg	agaaaggcaa	aactggcatt	agccgttact	75900
aataaaattg	cagcagcagc	agcaagggat	ggggaaaatt	catcagctaa	accgtcaaac	75960
tttggaata	gattggatga	agcaataaac	cctggagcac	ttttattacg	tagaggagga	76020
ggagtaagag	gaggacaaac	accccagagt	tcaatgctaa	caatgttccg	tctggagcaa	76080
actggtggca	atagtagttg	gtggactact	aatacacccc	ttattcaacg	cacaactagt	76140
gttggaata	atttagttgt	gcttgtaacc	aacctattgg	attcacaccc	tcctacattt	76200
aattaggaat	aaggaattat	acacatttat	ttgtatgttt	tacatgtatt	ttattcaaca	76260
ataaaggtta	tgatcataca	taatcaactg	gttggtttat	tacatagttt	ccattactga	76320
cccttccaaa	aatgaatttt	gtacctccga	atacaatttc	ttccctcggt	ttacaaagga	76380
aattgcacgg	agcgaatctg	ctcaccatt	caccaacaac	accctcatta	atcttcccat	76440
gagtgatgtg	aggattcagt	ctgttgctcag	cgagttcaac	atttctccct	attgcatctg	76500
aaatgtgttc	cattattcgt	ctacgcaaat	cgagtgcac	atctccaaca	catactttgg	76560
atacgagaat	gttggtgttc	atcccaaaca	actcaaaatt	atccccagtg	tcgattttga	76620
tagctggtgc	aaaacaactc	ccaaaatctg	ttttcatcag	agtattgtat	tcttccatat	76680
ttcgcaaaata	gaaaacagaa	atgggtcacat	gttgaggcaa	ctgagaaaat	gatttggaga	76740
cgtgtttctc	agataataat	tttcccgttt	ctgaatagaa	aggaagtaac	tcttcgctgt	76800
agagagcaaa	gaacattttc	atttcttcac	ctttaaaatc	aacttgattt	atagtgaaga	76860
ttctattgtt	gtgtttgatt	gaaaattggt	tcaccacttc	aaacatgcca	ttgatttggt	76920
tgtacatttg	cttgactgcg	catcttcttt	gcacctgtgg	ttgagtagat	gaaggggaga	76980
tttggtgccc	atttatctggg	ttggtatccat	tcaatatgtt	gtagccggca	caagggataa	77040
ctgggtccac	aaacttgaaa	gcctctctat	taaaccctct	ccctagagga	ataagtccac	77100
aacgaatttc	gtctgcgtct	tctgccttgt	ccattaacat	catattttaga	tgcaacaacat	77160
atctcagaat	cccattcaag	gtgtgttttt	ctggatagag	gcaaaataaca	tcattccacat	77220
attccccatc	tacataaaat	ccagtaacaa	caatctttta	cgcttttagat	aaaacaatag	77280
ttctctatg	agaaaatatt	gccttggtta	ttttggtgga	gagttgttcg	tcactgggaa	77340
tgggcgtaga	ggccaagagt	agagccttgg	aggaaaattt	cttgtccata	ggaagtctgt	77400
gccggatatt	gtctaccatg	ttagacatag	aagtattatc	actaaagtca	aagaagacta	77460
ctccattttc	tggagagggt	ttgttgaagg	atagatgttc	ttccatgggt	gtggtgggtg	77520
ttgtgatggt	ttgaatgagg	tgtggtacag	tcattatata	cacgtgggcc	ctttttaaaa	77580
gggggaggga	gggggagagg	gttgccatca	agcaatatct	tcgtccagaa	acacctggtc	77640
cagaaatggc	cataagatac	ttcctctatt	tctggagcta	ttacatttct	ggtgtaattc	77700
gacattggcc	cgaccagag	gtccaccctc	ggaacttgac	attcggtcga	gctagcggtc	77760
caccccctaa	actggagcgc	cctaaaaaat	ttttgaaaag	tttttgagat	ggaggaagag	77820
taaaattctc	tagtgaaaac	agaagggtat	accctctcat	ttctgggtcg	accagtctcc	77880
agaaacgcct	gttccagaaa	cacacaaaag	ttaattgtacg	tttctggagc	tattacattt	77940
ctggtgtaat	ctgacattgg	cccgaccag	cggtccaccc	tcggaacttg	acattcggtc	78000
gagctagcgg	tcacccccct	aaactggagt	gccctgaaaa	aatttttgaa	aagtttttga	78060
gatagaggaa	gagtaaaatt	ctctagtga	aacagaagg	tataccctct	catttctggg	78120
tcgaccagtc	tccagaaacg	cctgttccag	aaacacacaa	aagttaatgt	acgtttcttg	78180
aaccaacaat	ttctggtgta	atctgacatt	ggcccgaccc	agagggtccac	cctcggaact	78240
tgacattcgg	tcgagctagc	gggccacccc	ctaaactcga	gtgccctgaa	aaaatttctg	78300
aaaagttttt	gagattgaag	aggagtaaaa	aactcactat	atgaaagggtg	tgtagaacac	78360
cacatccttt	ctaggcacgg	ccatgtccag	aaacgcctgt	tcagaaaca	cgcaaaagtt	78420
aatgtacggt	tctggagcta	ttacatttct	ggtgtaattct	gacattggcc	cgaccagag	78480
gtccaccctc	cgaacttgac	attcggtcga	gctagcggtc	caccccctaa	actcgagcga	78540
ccctaaaaaa	gtttttgaga	gttgactaaga	tgagtagatc	gtaaaattct	ctagtgaaaa	78600
cagaagggtta	tacccacact	tttctggatg	ctcatagatc	cagaaacgta	tgttccagaa	78660
ataccataaa	gtccttttct	ctatttcttg	aaccaaccat	ttctggtgta	atctgacatt	78720
ggcccgaccc	agaggtccac	cctcggaact	tgacattcgg	tcgagctagc	ggtccacccc	78780

ctaaactcga	gcgaccctaa	aaaattttttg	aaaagtttttt	gagatggagg	aagagtaaaa	78840
tttcttgcc	gaaaggtgtc	cagaagtgtg	aacacacaca	cgcattctcg	cacctgtgca	78900
ccatcattaa	taaaaaatgt	cccctcccgc	tccccctca	aaatcgggtg	taaatagagc	78960
tcaccagcac	accacagaca	tcattctcaa	gacatttcaa	gtactgagaa	catcctctcc	79020
ttttgtgata	ttattcaaga	aaacctactg	aaatcgtccc	taaaatggat	tttgaaggaa	79080
ctaccagttc	tacccctca	aaaatgtccc	agttgtattc	atcagtgaag	aaagtgtcag	79140
agcatttcctt	tgccaatctt	catgacaagg	ctactcttgc	atcaaagggtt	attaaggacc	79200
tggaagggga	gaggaagaaa	atgtctaccc	caaagtcttc	ttctgatgga	caaaaactgg	79260
acaaggctat	gttgagcgat	attatcaacg	agtatcaggc	cgtaaagagc	actgcagata	79320
attccattga	atcgaccatc	aaggaaattg	aaaatgtact	tgaaagtgtg	cgcagaacca	79380
agattgaaaag	tgaagccaag	aacagtgtaa	cttcagccc	agaaaaagtg	tttctgtcg	79440
aggatttaga	aatctactcc	aaggggcgag	tgtgcaagg	tctcaagtta	aacgccaact	79500
gttcaagaat	tggaggcaag	tatgcagtgt	caatgagtat	caaaaaacac	aacgtctcct	79560
catttgagaa	caacaacaac	caagttttct	ctgaagaacc	cagggattgt	tttatgcttg	79620
aaacaaccta	tectcttggt	gggttcgaaa	cttctacaga	agatggaaat	acatatgcag	79680
ttttcttgac	tggtgttggt	ctagaaagat	ctctacctaa	atatgtacca	gtttctgaca	79740
tgaatgcagg	tattcaaacc	ctaaacatga	ctggtttgag	gatggccaag	cttctgttc	79800
tgtgcatggt	tggacgtaca	gaatatgaca	acttggaaag	tttttacatc	acttcaattg	79860
agacgcagtc	ttttgacgaa	gaggaatg	atgccagaat	gaggtgtcac	accgaagatt	79920
tggaagggaa	gaagcgcatg	aatgacgcac	cagcgattac	acctcatgtg	gccgtgtacg	79980
actacagtgg	gaagcggaaa	gaacaattgc	tctatatgat	aaccgagtat	gaaaacacgg	80040
ctagttggtg	caacgcaaac	gggtgtgtca	catctgacag	tggattttct	aacgaatgtg	80100
caattagtga	tatgaatgac	ttgtgctgtt	ttgtgactg	catcgatgtt	actgttaata	80160
atgaagaaca	tgaagaacgt	tctatgaata	ttgtggtcga	atctgacagg	cgtctttttg	80220
atgctagtcc	ttcccccatc	aagacggaag	aagatggaga	aaattcatcg	tcacgtctct	80280
cctctccaac	agttcctcct	cctacaccat	acgaaggtaa	cgcagttgtg	gagggggagg	80340
aagaagagga	agaaattgat	gaagacgaaa	gtagcaagta	tgaaggttca	gaagatgctc	80400
ttgttatgaa	gaaattagcc	aagctttcta	ctatgaaaca	aatgagaagg	gttaagaatg	80460
aacctgcact	caaaattact	tctgggggta	acaatagtag	cagtagtatc	aataacgaag	80520
atgatggtga	tgatgacgat	gccgttgacg	ctactgcatt	atgcccccaa	actgaagcta	80580
cagtgaaaaa	ttccttcattg	gccccaaacg	acgagagaac	tgaaaatatt	ttgtatgaaa	80640
ttatgcattg	ttctcttgct	aaaattttgta	ataatccatc	atctatgagc	agttaccgtg	80700
tattcaccaa	caaactccaa	gagtgtttga	ataccatgga	cgatagtatc	cgtcgccgtc	80760
caaccatttg	gactgaagaa	agtcaacaat	ttgctaaggg	tttgttgttt	gatgaggttg	80820
tcacatcaat	tgtggcacat	cagatggctc	aagatatttg	caagtctgaa	atatttgagg	80880
gaatgtttaa	cgccaactct	accaacatta	agggtaaata	tgaaggacaa	aagaagagtc	80940
tgtatgaaaa	caagcacatt	tcttctctgt	gcttcaaaaac	caacacggaa	tctaattgtg	81000
ataatgcact	atttgcgtgg	gtgaaatcga	aactccattc	tggcacagtc	atacctaacg	81060
tattctcctt	caaaatggca	tcagaaaagc	cctcaaaaat	gaagcgcaag	cgtacctcta	81120
gtgcttcac	atctaacgat	gaacaccaag	aaccatcaac	aaaaatgatg	aaaaatgatg	81180
aaggggaaaa	ggttgacaaa	gaatcatcat	ctccttcttc	atcgtctaca	ccagaacaac	81240
aacaacaagc	tggtcatgac	aaggaaaacta	tcaatttaac	tcccctcagt	ttcataaaaa	81300
tgccacgcag	taattgtcaat	ggctcggctt	catattttgc	tgaatatattc	ggtcaacgtc	81360
tttgtggact	gtctgatgct	tccagcacat	ttaagagaat	gtgcaagact	tttgaagatc	81420
ttgaaaatga	aatcatgagg	agctcattca	ctagactgac	tagatatgag	agggaggtaa	81480
ctcgcttgta	tgagaaatgc	aggtctcaag	ctgtagatat	tgaggaaaat	gaaatggatg	81540
ttttgtctca	ccaaggggaa	ttgtttgccg	agttcttgga	ggacccgatc	gcttactttg	81600
aagaagtact	ggagaatatt	aagagttgga	gcctagaaaa	cgtaaacacc	cctaagcgca	81660
aaaacaagta	tgcaaaaggtg	ctggtgagcg	ttaatgctat	tcgtaggaca	tatgaagaat	81720
accatgcgtt	tagcaagttt	gtaccaatgt	tctgtttcaa	tctgattaag	agagaattgg	81780
aaggagacaa	ctatacccat	gacgttcact	tttcttcac	ttgcctgtgg	tacctgactg	81840
taatgaccag	gaacaggatt	tgcgatgtgc	tccagtacat	caacaacaac	aataatgata	81900
acgaagaaac	cgatattgtg	gaggaagagg	aggaaggaga	aggagaggag	gataaaatgg	81960
aagaaagtat	ggacgtagaa	caacagaagc	aagttcgcaa	gggaggggaga	aagaagggtc	82020
aaaaattcaa	cagtattggg	gatcaagtca	ttagaaaatt	tgtgaaaagt	ttgtgtgaaa	82080
attcgatggt	agtttctatt	gcaattaata	gtttgatctc	tggaataagc	tggaataagc	82140
agaaaatccc	tccgggttct	ttgaaggatt	ctagcacaat	caataccctt	gatgaggtct	82200
caaggtttgt	gttttagcgt	gtcaaaaatc	ataggaaaat	caatggaaca	gatgataaat	82260
atgaaactgt	ttttggagtc	gtacgcgtg	tggattcaca	tattgtaggc	ccctttagta	82320
tacctgttga	tttttcaagc	gcaggactag	ataaggcctc	atgtggcaaa	ttgtacgtta	82380
acaccataga	cggaaagggc	attttgacaa	tttcacccaa	atatgattca	ttaaacgatg	82440

aggatgttga	ttctactaca	acagacaagc	tagagaagga	tattttgcac	ttgtctaagc	82500
atgacacctt	tttcaatatt	aataagaata	aggttcttcc	attctataat	atttctccta	82560
gctcttctct	cactgaaaag	aaaaagacaa	aattcaatag	gaagaagatc	tcatctggta	82620
tgagcaataa	taatggcatg	tgtgtacaaa	ctccttctag	ttcaaattca	gtctcttccg	82680
tctcgtctat	tgtagctcct	tcaccttctg	tcttggtctc	atcttgctcc	ctttcttcta	82740
caaagaaaaa	gagcatctgg	aacgagaaca	tgtttttgac	atctaggaac	atgtggagg	82800
gtggatttgt	tgtaccaccc	aaactttgca	gttttattgt	taaccataga	cacgctgtaa	82860
aacttgtagc	tgaactgca	cctaaaacaa	agttgtgtag	gaatattatt	gataggaata	82920
ggaagattag	atttaacggt	ctaaagaagg	tatgcaagag	tgttagcgcc	tttaccggcg	82980
agtctacata	tttgctcaat	aagaatatga	ctgcaacttc	acctagtgat	ttgaacctat	83040
gtattttatac	ttcatcttta	aatgacccat	tgtatacttg	taaattgacc	catgaagagt	83100
accaggacgg	aaatgctttg	gatgactatg	gtgcagtttt	tgtaaactat	acattttaaga	83160
gtatcaaatc	ctgctcgtcc	aaagatgaaa	ccgctgacga	caatgctgct	gctgctgatg	83220
atgatggttc	tactacttct	acctcttctt	caactgatac	tgatgctgct	gcaatccaag	83280
actttatgca	tgtaattgatt	aagaaaattg	atgccatgaa	agacattaga	ggaaagtaca	83340
agaaatcttt	ggcaaagaaa	acaaagaagc	attaaatgta	tacctttatt	tttaccaata	83400
aaaatttat	gttttcatg	ttgtgtttta	ttcaaattca	tatggttcat	ccccatttaa	83460
atcgtctaac	cccatatccc	agtcaaacgc	tgtgctgct	gctgctggtt	caatgtctga	83520
gccatagatg	gattcattga	catctaaaag	ttcatctcta	aaatttctcg	gtgtacgcgg	83580
ttgacgaata	attggtgcaa	taaatgcctc	ctcttcttct	tcttcttctc	ctctgactcc	83640
cgtttccttg	acagcttgac	gttgcgcttt	tattatctct	tcaacgttcc	tccctggagt	83700
agcgggagg	agatcaaaa	caatactgcc	aagctcagag	tctatatcag	taaaggacat	83760
atcagtagcg	ttttcttctt	ctgctcttgc	tgtagacgga	ggaggaagag	gagtaggctt	83820
gggtagaact	ggtgtagggtg	gagggttagg	ggtagtagaa	cgagagtcaa	atcttacacg	83880
ggaaataggt	ggagaaggag	cagtattagg	tataggagga	ggaggagggg	gactagaaat	83940
gagttgaacc	ggaaagggct	gctgctgctg	ctgctgggga	ggaggaatat	tactgggtga	84000
ttgtttgtga	ggagtagggg	gaggaggaat	attattgggc	ggctgtggtg	gttgctgctg	84060
ctgctgctgc	tgtctgcagct	gaggcgccgg	ctgtgattgg	ttgcctaaaa	agtacagcgg	84120
gtactgttgt	tgtaatgtgt	gaggagcata	tgggttataa	tactggttat	attgataata	84180
aggatttaag	tatccttggt	aagtaggttg	ttgaacctgg	aaaggtgcct	gtgggtggagg	84240
aagaggaaa	ggaggagggt	ggaaaggggg	aaattgttgt	tgttgttgct	gacgttgctg	84300
ttgttgttgt	ttgggcacaa	acactccagt	cgataaaatta	aatgcttgac	ttgtttcctc	84360
cactgcactc	tttatcattt	tttcaaccgg	tcttgctgaa	gctcttgcta	aaacaccagc	84420
tgcacatttc	tccagcatgg	agatagaatg	ctctgattta	ctcctcatcc	ctcctccaac	84480
caaaagatcc	atagtcgatg	acatttttac	ccacacgcta	ttattttatag	ctctaaaata	84540
acgaggtatt	gataggccac	gtacagtggc	gttagagtta	tcagaagcta	cagatacaaa	84600
attctttatca	aagtctttgt	tcggtaatat	acaggtacca	aacaggttat	tacacacagt	84660
agcaagtgtg	tcaaagtata	ctcgttcttt	cacagtctgc	tgtatacgat	cctccacgat	84720
actgctcatt	aaattgttta	tctctgttgc	atttccatcg	ttaagagatt	tcaatatttt	84780
tgaacttaat	tcttgccact	gaatctctcc	tgaccgtata	tctttcctat	gggaaattgc	84840
ccaacaacta	ttagccagaa	gagctttcac	caattcagca	cccacatttg	tgacaacaaa	84900
agtgttttgt	ctttttctct	cgaacttttt	cttcccttgt	tcagatgaca	ataaagatga	84960
tgaagtagga	aaagccttgg	ttttatttgg	ctttaaaagg	ttagctaaaa	gagaccttat	85020
tctgtactgt	atagccttct	ccttgaccac	tgccctgctc	gccaagtatg	acacgtaatc	85080
gtattgtggc	gcattttcac	tccttacacc	taaaatagag	cccgaagtta	tagcacctct	85140
caatggtcga	gacacggtea	tcatttttcc	agcaggcaat	gctaattgct	ttgattctgt	85200
gtcacttctc	tttctttttt	tagagggagg	aaatatgtgc	tttttcattt	gttttgcccc	85260
cttcgataca	gatgtagtag	gtgtaacatg	ttcgtcttca	acattctcct	cttcttctct	85320
ctctccatct	ttttcttctt	tctcttccat	ctgttcttcc	ttttcttccc	gctgctgttc	85380
ctccccttcc	ccttctctct	tccttctctc	ttcttcttca	ttacctctct	ctcttccacc	85440
accttcttga	tcagtgtctc	tttttccctc	atcattttct	aaatctgaac	aatcgccagc	85500
aacacaatcg	ccactccttg	acagaatttt	actagtacga	atcttacgcc	tttgccattt	85560
cttactagaa	gttggcgctg	ccaattttag	gacctctcta	cccatgtttc	gttggaatct	85620
ggattcgcg	tcactcctct	tttttcgttg	ttgggttcta	attatatgct	tttttaaagc	85680
agtattaggt	ttatgttttag	gtggcatatt	tgtaaagggtg	gataaacatt	acaagaaaac	85740
aaaaatatct	aagacattat	tttattttgt	aggcaccgct	gttttagtat	gtacatcaaa	85800
ttcaaaaacca	tcttccatta	catatatctc	aatctcttct	tcttctctct	ctcctctttg	85860
tttggacaaa	taactatcac	tacattttct	taaagctaga	caacatagag	tacaattata	85920
caccttttct	aaagcagcct	ttgtagactt	acaatcagaa	tacaatagtt	tatgggtcaat	85980
tagtatatgg	tcagggttag	ttatttcatc	ttcacatata	caacaaaaaa	ctgtagggtta	86040
cccatctctt	ctaccgctac	aatcataaca	agtcagtgtta	atacatccac	ctaattttat	86100

gttaggtgtc	ttacatttag	ggcattttac	agtcttaatt	tcctgtcctt	ctttctcaac	86160
cttccattct	aatgtcctct	tttcatccac	agtgaacact	ggatccttta	ttgattcaga	86220
tgtcatgttt	tgacgtacaa	tattttacct	tttaacaatt	tccatttttg	caccttgggt	86280
aggtctgata	cccgataagg	cacgcctatc	tacacaatga	aggacaaaag	aatgggctct	86340
atggaacaca	cattctggcc	gtccaacac	catttccctt	ataagattct	tgacaaacat	86400
tttcgcccac	aatcctaata	cactctcatg	gttcaatttt	tcatagacat	aaaaataggt	86460
tgcaattaga	ggtatactca	cacctacaac	ttgttcaaaa	tctctatcat	caaaaccccc	86520
aaaaacactg	tttaagtaat	ggactttaac	cctattctga	aacctctccc	tggtgtttgt	86580
gattacagat	atgtcgttga	gaagccccct	catattctct	ggcatacga	cagtcgagtg	86640
ccagttggct	agcgctagaa	gaaccttgcc	ctccggagac	tctgggttcc	atcctcgaat	86700
cttcttttcg	ggagttcttt	ttagtataga	aacagcctgc	ctacaatttg	tcacctcttc	86760
atltggaggt	acaaaatttg	acgccaagaa	accgcaggac	aatctcatgg	gccatttatg	86820
gagtcgagg	ccagtatcat	atcttactag	ctcatcggtg	ggagggtcat	caaattcagc	86880
atcttctaca	tcctggaagt	aatagtcttc	cagttgttcg	ggaggagggt	acggcaatac	86940
tggaggaggt	acatccactg	ttacaatgtc	ttccatctca	ttaggctggt	cacatacatt	87000
gggtagtaaa	cactgggtac	agatcaggga	acatttgcta	tcaccagttt	ctccctccac	87060
catcttaaa	agtttaacgg	gcggctctga	tctattatcc	accacagtaa	atltttgaatc	87120
cttgtctagt	gcaaagtctt	tgtccatctt	acacatctcc	atgaattcgg	cctgaagaag	87180
ttccacagta	tagaagcatt	tctctgtctg	cttgatgagt	tcggtgatga	gaagtttgtt	87240
ggctgcaatg	tattctgggt	cgacattttc	tcctgtctaa	atcgccctct	cgaataaaga	87300
tgaatttcg	tgacatctac	ctataaacctg	taccatata	tctccacaac	atgggcacct	87360
aaaaacgtcc	tttagttttt	cactaccagc	cgagaccttt	tcgtacacat	acattaggtg	87420
ttttatagta	caatcagggc	atagttttgc	agaacacgtt	tcttgtctg	cttcaaaagt	87480
gtccatgaaa	gcattgtacg	ggaaataatc	gtaagaggct	ttcatgtttt	catgagaaca	87540
gttgaagaat	tttggtttca	tgtccaaatc	cgcttagcac	aacctcacat	ttttatttgt	87600
ggtattccca	ctcaatatct	ttgtaacgtg	tgcttgtatt	acctccactt	cgtcagatgg	87660
ttggtacacc	atcgccgctt	gactttctcc	ccaagactct	tcttgcaagt	tctggctaag	87720
catgcatttg	gogattgctt	cttctagatg	catattttca	gcgtttttcat	gttcgaggat	87780
gaaagatcct	agttttctgg	ctgaagggtat	gccgtgtcta	gataccaggg	acagtgttgt	87840
atltttctaag	agcgaagcgt	tcggggagat	tattagacca	gaaggcgaca	ttacggattt	87900
ctttggctcc	cagtatctgc	tcgtatctgc	tgccaaaaga	tagggaacat	tgtagataga	87960
gtatctgtac	actatatcaa	tcattctggc	tatttttgtt	ccaaaaagaa	tatggtgtaa	88020
ttcttggggc	acatatccca	tagcgtcata	cacgggcctg	gcagcgtaca	tgcaagatg	88080
acccttattg	tagtggttgt	tgggtgatgg	accacttaaa	aatgcttctg	ttacggagaa	88140
ggccataatt	ttccaaatcc	tgtcatttgt	tcttccttca	tgcatatcag	aaccaataaa	88200
aacactaaac	gttctagcgt	ctcttatttt	tgtccccaat	tgcaaaagcag	agaagggccca	88260
catgtaccac	tccttatgaa	acacatcctt	caaagtgttg	agagaaacta	ccacatcttc	88320
ataaatgaca	cttttcatgt	ggcgggtgctt	cttgttctga	tctcgtttga	tggcaacaag	88380
gagtaagaaa	atgcgtatca	aatccatatc	tgatatcatc	tcggcgtacg	cctcacacaa	88440
cgcattcttc	tcctccactg	catcgccctg	atttccacgt	ataatttcat	ccaacatcaa	88500
gcgcgtatlt	ttcatgacaa	acatgccact	ataaaatcga	tgaattagat	ccttgtagtc	88560
tggcccatct	tgttttttgc	aactctttaa	atgtccaatg	aattctttgtt	cgttttggtc	88620
aagagttgta	aatggaactg	ggctcgtcgtt	acttttcatc	ttgatgttca	tgtcaggggc	88680
tttcttcatt	aatagagata	tctctttcaa	aagaccggtt	ggattaatca	caactgtttc	88740
tccttccaaa	agttcatccc	tataagaatt	gcggttgacg	atgtccagta	tggcctccag	88800
aagatcacca	ctttctccca	ctatacccac	aattcgttca	taggtggacc	tcttattatt	88860
tctgcctccc	attaggtcgg	tgatcaagat	tgcaacttct	gggacagcag	ctgcaactat	88920
ggatgttaca	ccttccatgt	tgaacgaggg	aatgaatgag	ataaagtctg	gggcctccat	88980
cttatatacc	cttcagatgt	gaaaaaaaac	aacatgagga	aaaaagtaaa	aaattcaata	89040
tacatgtatg	tctttatttg	ctaggaagaa	cacatttcac	aagcaccagg	gttattgata	89100
gaacataaaa	gagcagcagc	aggatctgga	acaacaatag	aggacttctt	ctcctcttct	89160
tcttctcgga	caggagcaga	aaaagctgca	acaggagagg	gagcttccct	cttgacttct	89220
tggagtacat	tcttgctgac	tgtgaactgg	acagctctag	ctgcgccctt	tgtgcgtaga	89280
taatagagag	tcttgatccc	cttctcccat	gcatacatag	tcattgaccg	caccttgctg	89340
agttcggggt	cttccacaaa	caagttgagg	gattgagctt	ggccaacaaa	catacctctc	89400
tgaatagcca	tgtccaaaag	agtacgagga	ttaatttccc	atacagtttt	gaatagtccc	89460
ttggtcgatt	tagggatatt	aggaagcgtc	tgaatagatc	caccacttgc	cataatcctc	89520
tgttttagtta	ctgaattcca	ttctcccagt	tttatgagct	ctctaataac	atattcggtc	89580
accacttgga	atgatcctga	aagtacatta	cgattataca	tgttgagcgt	taaaggctca	89640
aaggattcag	agtgcccgag	gatctgtgca	gtggatgcag	taggcagggg	agcaacaaac	89700
attgaattgt	gaacaccata	cttcataatg	tcccttctca	attgctccca	atcgtgaatt	89760

ggcaaagagt	tgaatatata	gtccctattc	ttaatatattt	ccttccccat	gtcaaattgg	89820
aaaatccctt	tgctcaaagg	actaccctca	aacagctcat	atgtttctcc	cttttctttg	89880
gcaatttcac	atgaagcttc	caaggcacca	tagtatatag	tttcaaaaat	cctcttggtta	89940
attagtccc	cttcttcaga	ttcgaagggg	attctgagtt	tgaagaacaa	atctgtctagt	90000
ccctgcacac	ccaatcccat	tgccctagtt	ttcatattgg	aatgcccggg	cttgtcaact	90060
gcatagaaat	tgacatcaat	caccttgctg	agatttctgg	tcatgatttt	tacaaccttc	90120
ttcatctccc	ggtaatcaac	atagggcctt	agggagggga	tgggagaata	cttcacaaac	90180
ttgttgactg	cgatagaagc	caaattgcac	actgcagttt	cctccgaatc	actgtactgg	90240
acaatttcag	tgacacaaat	tgaagacttg	atgatgccga	cattttcttg	gttagacttt	90300
ctattgatgg	tatccttaaa	gcacacataa	gggtgtccag	tttcgatacg	tgacagaatta	90360
atttggtcga	ataatgcacg	tgccctcacc	acctctttac	ctttgccttc	agcctcatat	90420
ttctcgtaca	acgccttaaa	ctcttcgcca	tggacgtcgg	aaaggccagg	gcactcgtgg	90480
gggcacatca	gggaccaatt	cttcccagcc	ttcactctct	ccatgaagag	atcagatacc	90540
cagatagctg	ggaaaagatc	cctcgtcctc	aaatcttcat	taccggcatt	ctttctgcag	90600
tcaataaagt	ccttcacgct	cagatgccaa	tcagaaatat	agatggctgc	agctcctctc	90660
ctcttgctc	ctccctggct	aacctttttc	acagagacgt	taaagatttg	gagggaacgc	90720
atgagaccag	ggtgggtacc	actccatgac	gaaatggggc	ttccttttgc	cctcaaatca	90780
tgaaagtgga	tgccgagtc	tccagcagtc	ttggagatga	ttgccgcctc	cttaagagta	90840
tcataaatac	cctcaatgct	atcatcttga	aggcccagaa	ggaagcacga	ggaaagtgg	90900
ggtgtgactg	ttccacaatt	aaacagtggt	ggggaagcat	gggtaaaata	gtgcctcgac	90960
atgagatcat	atgtttcaat	gacagatttg	atgtctgac	cgtgaatgcc	gacagccaca	91020
cgcataatca	tgtcctgagg	gcgctcaacc	aagattctct	tctttctatc	agtgggggaa	91080
ccaatcttga	tcaataaaga	gtattctagc	gtcctcagtc	cgaaacaggt	gaagagataa	91140
tccattttgt	aatcgataac	agcatcgaga	atttcagcat	tggccataac	attttcatag	91200
taggtatcat	taactactga	tgctggttta	ccagttccag	ggtggattgc	gtgcctcaat	91260
ttctgagtgt	ttgcactaaa	actattccac	cttttggtgt	ttttgtggat	gttcgaacag	91320
atgaatcttc	ctgccagttt	tccaaaatca	gggtgggtcaa	caatttttgt	ctttgcataa	91380
tcggccagaa	aatcgctccat	ttcttggaag	gagatggtag	cggaagacg	gtccatgatg	91440
tgagatgcaa	gttcttgagg	gttaattgct	ttcttggtcaa	gcttgggcac	atactggttg	91500
actggaagac	aggcattttc	aatcctcttg	attatctttt	caagactaat	ttcttgctta	91560
gtgccattcc	tctttgagat	gaatgattgt	tgctggttag	aaccattttt	atgttattaa	91620
attgggatgt	tggtatttgg	agtattttcc	tatacgtctt	ctcggctcga	ggagggtgaa	91680
agatgtgagt	gaagaggctt	ggacggtagc	tgtatttata	ccggcccggg	gaaggctcgt	91740
ggatattttt	gaaggagggg	aagaaatata	tagtattata	cgtagttctt	tgatgtagaa	91800
agaaatcaga	aggtgaggtg	acagctgctg	ctgctgcagt	attagtagtg	cagggtgcaa	91860
actgttcttc	aaaaactttt	taaaaatttt	tctggatcac	tcgagtttag	ggggtggacc	91920
gctgtctagg	cctgactgtca	agttccgagg	gtggacccgt	gggtcgaccc	aatgtcagat	91980
tgaccagaaa	ccgactgttg	ctccagaaat	gtgcattaac	ctttgtgtgt	ttctggaaca	92040
gacgtttctg	ggcatagctg	cgccagaaaa	gggaaggtgc	aacatactgt	tttcgctata	92100
gaattttact	cttcctcact	ctcacaaaat	tttgaaaaat	ttttctgggt	cactccagtt	92160
tagggggtgg	accgctgact	aggcctgatg	tcaagttcgg	aggggtggacc	gctgggtcgg	92220
gccaatgtca	gattgcacca	gaaacgactg	ttgctccaga	aacgtacatt	aacctttgcg	92280
tgtttctgga	acaggcgttt	ctgggcatag	ctgcgccag	aaagggaagg	ctgtatctta	92340
ctgtttctgc	tagagaattt	tactctttct	ccctctcaca	aaattttgaa	aaatttttct	92400
gggtcactcc	agtttagggg	gtggaccgct	gggtcgaccc	aatgtcaagt	tcggagggtg	92460
gaccgctggg	tcggggccaat	gtcagattgc	accagaaacg	gaatagctcc	agaaatgtgc	92520
attaactttt	gtgtgtttct	ggacaagtgc	tttctggtct	aggaggcaac	tagaaatata	92580
agctttagc	ctagaattta	gtacaccaag	gataaagaaa	tataacaaat	accatactat	92640
gttttatttg	agtaatatag	tatccaagtt	aaaataacaa	tcacgcactt	gtacattaat	92700
agtctgtaaa	agatgggcaa	acgtacacct	gagagtattg	tcccagacaat	cgtacatgaa	92760
ccctctagtg	cacgtatcac	gttttcgccc	tattaccatt	gatgcacaaa	tttctctctt	92820
cattctaata	gaggcagatt	gtttgtcaaa	ataacactcc	ctatagtaac	aaccaggatt	92880
tccattggat	tgacactgta	gttgttcgaa	cggttttcct	gtataacatt	tcccagctgc	92940
gtggtagtga	cgacacaaac	tcatgcaaga	atgtgcgtgg	gtagtttcat	ttctaagaat	93000
ctcccacgca	ctagaggcac	agaaaaaatc	aactgctgcg	tggcagtgac	gcacggccac	93060
acacccgcat	tcttcacacc	ataaaaggac	atgatctgtg	ttagtcgtca	catctctcag	93120
aaacccccat	gtgcattgac	gtcatgggga	attgctattg	caccacctta	tggtgtataa	93180
aagtggccct	gggtggcacc	ggttagtaga	cagaaaacaa	cgtcaagat	ggtgctcgtc	93240
attacccacc	tctctctgtt	gttcgtctgt	gctgtagtag	cttccgtcgt	ttttacaact	93300
gaaggagcta	gtgtgagagt	gaaacggtgt	gctgttagcc	cgtgccccga	cgttattgac	93360
cccgaccacc	gctgccaagg	gcgactgtgc	cgcagggtcta	ctcgaggagg	tgacgacgac	93420

gacgacgatg	acgatggagg	aactttcgat	acagtagggt	ctggtatact	tggacgcaaa	93480
aagcgtgccg	cacctccacc	tgaggatgaa	gaagaggatg	atctctaccg	caaaaagcgt	93540
gccgcacctc	cacctgagga	tgaagaagag	gatgatttct	accgcaaaaa	gcgtgccgca	93600
cctccacctg	aggatgaaga	agaggatgag	ttctaccgca	aaaagcgtgc	cgcacctcca	93660
ctcggagatg	aagaagaagga	tgagttctac	cgcaaaaagc	gtgccgcacc	tcacctgag	93720
gatgaagaag	aggatgagtt	ctaccgcaaa	aagcgtgccg	cacctccacc	tgaggatgaa	93780
gaagaggatg	agttctaccg	caaaaagcgt	gccgcacctc	cacctgagga	tgaagaagag	93840
gatgagttct	accgcaaaaa	gcgtgccgca	cctccacctg	aggatgaaga	agaggatgag	93900
ttctaccgca	aaaagcgtgc	cgcacctcca	cctgaggatg	aagaagaagga	tgattttctac	93960
cgcaaaaagc	gtgccgcacc	tcacctgag	gatgaagaag	aggatgattt	ctaccgcaaa	94020
aagcgtgccg	cacctccacc	tgaggatgaa	gaagaggatg	atctctaccg	caaaaagcgt	94080
gccgcacctc	cacctgagga	tgaagaagag	gatgatttct	accgcaaaaa	gcgttaaaact	94140
acgcacgaaa	gtgacggtgg	ttgaagaata	gactaatatt	gttgatatgt	taaccccttt	94200
ttttcatgaa	atgtgtacac	acctgctata	tatacgtgca	tatttgaata	aggaataaag	94260
tttatctgca	tctgtatttt	tttattagta	tcctttttgt	gggataaata	atcggattatt	94320
gtgtccaaag	atagtgggtg	catgatgcc	tatacaaagc	acgtgatact	ttccaagaag	94380
cagagtcgta	tgaccattga	cgtcaatgcg	tgggcagagg	cggagggtgt	gataaagcgt	94440
ttctgagaaa	cattgggcgt	atgacgtcaa	ctacattatt	cttccctctc	ctcctcctat	94500
tgccctctgc	agttcagtat	tttatttttc	ttcctataca	ataaaaaagta	tcagatgaat	94560
atttttactg	ttttctgtt	catccctctt	tcctattgta	aaaaaaccaa	taactaacta	94620
atcatggata	acttgaaagg	ggaatttgtt	gcgcttaaaa	cagacctcac	ccattacaaa	94680
acacagttgg	atagatctat	attgggtattt	gttgatgttg	ttggtagatt	atatgttata	94740
gtaaatagtg	aacaacacg	taaaaaggaa	ggtctagcaa	ctagatgggc	aaagcaagcc	94800
atcagatatac	aacaattcaa	ggacgaataa	aacaacaaat	ataatgctct	aacaataact	94860
ttggatgata	tcacttcat	ttttgatcat	ggaggagggt	tcaaagagtc	aaaacataag	94920
gccataattg	aagcgaggga	atactctaaa	ccgctgaggg	aattagatgt	catgtttacg	94980
cgtatagcgg	acatgttaac	cttgactttt	atgactgtgt	acaccaatat	cattactgaa	95040
tttagacact	ctagtgaaca	agccactaat	agtataaatg	tcaccctcgg	acgtcttttc	95100
ttgtgtgacg	acttgtgcaa	tcaattacca	aaagaagagg	aagaagagga	agatttgaaa	95160
cagaaattca	ttactttcca	tcgcaacct	tacatgctgg	acacacgcct	aaagaagaat	95220
ttgataattt	tcaaagatgt	catacaacaa	cttcacgtga	ttttgcaaaa	ggataacctat	95280
gctgtaaaag	aagggtgtgc	cattagatgt	gcgaaacaga	tgaacgaaat	aagtcaatac	95340
aggggcaacc	tcaaggataa	ttacaataca	ttttcaaaca	ttttgaatga	aattgtctac	95400
atttttgatc	acgggggaca	ttttgaagaa	gtaaaaacaca	aagccataac	tctgactaga	95460
aattacttga	aaacactcat	gggattataaa	tgcattgttca	aacgcatact	cgaaatgttg	95520
tcattgactt	ttctaacagt	gtacactaat	gttatagcag	aatttataaa	cgctagcaat	95580
atctctgata	gagagatcaa	taattatctt	gtccaacttg	taacatgtaa	cgaattgtgc	95640
aaccaactcc	ccaaacctaa	acaataccgt	ccctcagtt	tgatagataa	catagcttat	95700
ttttctcttt	ctgtccaaaa	acatctgagt	gggtttcttt	agtatgtgga	gtattattgt	95760
cttgaagtc	ataagtattc	aacctctcct	actggttaaca	tctcttctc	tctataatcc	95820
caatatggat	agttgttgcc	tgatatcgag	gataacacca	gaactggctg	ccaagttaac	95880
ctggatcttc	ataccagaaa	acaacttcaa	gattgtccag	aactcactcc	cagacgacca	95940
agttatctcc	caattcagat	atctcgacca	tagacattgc	tatacgttta	tggagatttt	96000
gatggcaaac	attaaaaatc	aagacaggaa	acaaaaacac	acagccatat	gtgaattgac	96060
aactggaaga	gaaggacttt	tatgtagaag	aaccatacct	gtaatttttg	gttcagagga	96120
aaaacagaga	gagttattgg	ggaattctcc	tgaagggtgca	gaaattttca	ggcctagaga	96180
agttatgcaa	gtaattggta	ctctcttgga	caagaaacta	gaaattgacg	acggtatagc	96240
ttctgtaaa	gctgcctct	gtgctggttc	atcatcgta	tacctaatca	tgagccacat	96300
agtgaataatg	accttttctg	atacacaaa	catcaaggat	ataaacgaag	aatatttctg	96360
agcttttata	tttctgtata	caaatctcct	caacctgaa	ttcttcaag	accttatctc	96420
tttgctcaag	aatttcagga	aggaacatgt	tgcccatcta	gtaagacgtc	tagaacactt	96480
tctcatgcta	tggacccttt	ccaagatgag	gttcacagaa	atggaagaaa	actacttccc	96540
aatctccagc	gatagtgatt	acggcatctg	tgaaaaatgt	gcacgaaaaa	ctcccaataa	96600
caagctccgt	atttttaggg	aacgaaaatg	ctgcgataga	tgttgccctg	tttatcaca	96660
acaaccgcct	cgggaggtgt	ataattggga	tggaaaaata	acccaacaat	ccaataaagg	96720
ctacattaat	gcaggcgatg	aaattatcgg	catgctaaac	tcaaatgata	agggaaaaac	96780
attccctcct	atacctaaga	tggttgttacg	aagagtggtg	gacggtgtct	acgggcaagg	96840
aactatcctg	tcaaagattt	tgaagttcag	gcaggcaaat			

gaacacgtta	gaaaaattct	gttcctggga	aagattcaat	actgaagttt	tgctcccttg	97140
gcttggttac	actattgagt	ctaagtggca	gaactgggaa	tcttttctgg	gttattcgag	97200
taccagatat	aagggaactgt	gggcctttgt	gaacaaacag	gaaatatctt	ccatgaaaga	97260
ctcctacata	aaaattgaag	acatcgacca	gttattgagg	agtatcttgc	aagaccagaa	97320
gggtgtattt	gagaccgtct	gcaaaaataa	gagcagagat	ggtttgtgaa	ttggccacac	97380
tgattccgat	accgactaaa	aaggctgatt	gatggcaacc	ccccccctc	cagactcagc	97440
cgcattgagta	taaatatggc	catttctcac	accacagcat	cattccctcg	tcatcggtcc	97500
taccgtcaac	ttccattatt	actccaataa	taccaacaac	cccagaaatg	gagtcaatca	97560
aactgttcac	cgttgctggt	ctgaatatgg	agcaagccaa	ccaagtggct	gaagaaatca	97620
agtcagaata	taaaaccgag	gaggaataa	ggattgcccc	ggaagtgttt	gacaaattca	97680
ccaaaaaact	cattatgcaa	gtagatacgt	ctaaacactt	acttacaaga	gaaaaaccca	97740
accgttttgt	atcccccccc	attgtccatg	aagatctctg	ggaaatgtac	aaaaaagagg	97800
ttgcctgttt	ttggacattg	gaagagattg	atttcgaaaag	ggatcctaaa	gattggggaga	97860
aactcactca	agatgagaag	gatttcattc	tccagattct	ggcgttcttt	gcacccctctg	97920
acggaattgt	aattgaaaat	cttacaacac	gtcttcgtca	agtggcgag	attccagaag	97980
cgaggagttt	ctttgacttc	caagttggaa	tggagagtat	tcatggcaac	gtctacggag	98040
aactgattga	tagactgggtg	cccgacgaaa	aagacaaggc	tatcttgttt	aacgctgcac	98100
aacacttccc	cgccatcaag	aagaaggagc	agtgggctat	taattggatg	caaagcaata	98160
acgatttggc	ggaactaatt	gttgcccttg	ctgcagttga	aggaatcttc	tttagtggtg	98220
cattcgcatc	cattttctgg	atcaagaaca	gggttatctt	gcctggcttc	acctcctcca	98280
atgagttcat	ttctagggac	gaaggtcttc	atcgcgactt	tgcatgcatg	ctgttgaaaa	98340
agggttttgt	tgatacccca	tcaagagaaa	ggattcttga	aattgtcact	gaagccgtcc	98400
gaattgaaca	agaatttctc	acagtttccc	tgctgttaa	attagtggga	atgaactgca	98460
agttgatgag	ccagtacatt	gaatttgtgg	cagataaact	attggttgaa	atgggactag	98520
aaaagcacta	taattgttacc	aaccctcttc	cattcatgga	caatatttcc	ctcgagaata	98580
agaccaactt	ttttgaaaag	agagtcgccc	agatcaacg	tgcccaggtc	atggcttcta	98640
tcaataagat	caagaaggac	caacaaaccc	aagaaactgg	ttctcctctc	ccaattctga	98700
ctgcacctcc	tccagtctct	tcctcatcat	ccgaacaaga	agatgttgaa	gacggcgctc	98760
gggactacat	cagttatgac	gatttttagt	tcactatttg	tgtcaatagg	ttgtgtattg	98820
tattattatt	gttataatat	ttttaaaaaa	taaatgttct	ataagactaa	aacaatgaat	98880
ttacttccaa	tattcctgac	aacctttttt	gttgcggtag	atgcatgctc	ttgctctacc	98940
atctgccttt	tacctgatgg	gaagaaacaa	cccttggttt	ttgattctgt	attagaagag	99000
gtggtatacc	ctacagatgt	gtgtgggcca	aagggagctg	gcgaattatt	cactggtgtg	99060
gatcttttga	ccctctgtat	aggaggtaaa	aacaatggag	gtgaatggtc	aggaaaagg	99120
ccttgtccaa	ggatcaataa	cgctgtcggt	gaacgagatt	actcccttga	cgaggaggat	99180
tgtaaaagggt	ttagaaaggg	gttccgaatt	cctggcactg	accattttca	tactgtcttt	99240
tccttttgtt	gggtagacag	agatatgcac	gccaagtggg	tgcgcaacaa	aataaacctt	99300
ggatatagtaa	ctgatgatga	agatttggta	gattctggta	ttaggactaa	attttaaatac	99360
tcttctaaaa	tttttggtaa	aggattcaat	ccgagacctc	tttactccct	cgactatcaa	99420
gagaggatta	agatattaaa	gtctcatttt	aacaagagga	cgggtaattt	ctttgctcga	99480
ggccacttgg	ctccggtctg	agattttttc	ctcgcttcag	agagatgggc	aacttttgtc	99540
ctagagaatg	cagtacctca	aatacagaac	cataacaatg	gtgaatggaa	agatattgaa	99600
aatcgtgcaa	gaactacgcc	aggtgccgcg	tgggctgaga	ctggaccaat	attttaccaa	99660
cacaagaaga	aggaatatct	agacaagaag	aagaagtaca	tccttatccc	tcatgccctc	99720
tacaagattg	tgtacgacaa	gaataacaag	gaattgttcc	gtgtacagag	tgatatgtct	99780
tggaataaaa	atacataatt	aaagaattta	tattgtttta	ttttgtcatt	tatttgatac	99840
atttgttgta	tgttttaaag	acaatacaga	taattttcct	ccatgacgaa	acactgggtga	99900
tatgtctgga	gctctgatga	atgtcttgcc	agatgccatc	ctagtttttc	tcctggggaa	99960
ggtgacgtcc	gtttgagaga	gggttcgctg	ggtttcaact	gcagccaaga	aaatatccct	100020
ttgtccagtg	acggtaaaatt	gtgtaccgtt	atttagtgtg	atgacattat	ccctctctga	100080
ttgtttcaaa	gcttcataat	cttctctctc	atcttcatag	tcatctgcaa	cattctcagt	100140
gtctaataga	acaataacag	agttggggga	aaatcagtc	gtgctacgtc	cagactggcc	100200
agtattttta	ctgctccatt	gtgtaggatc	atgctttcta	ggcttgtctg	aatgtagacc	100260
attatagtgt	gctctcccca	tcataatcata	cagtgtacac	actagtgtat	taccaatgcg	100320
ttttccactc	ttattgttat	cacaataatc	ctcatcttct	ccaacaagag	ttgagtacac	100380
actggaatct	gttagaattg	tatgagaaac	taatgcctct	ttaaaaatat	cgctactctc	100440
ttctaaccag	aaatctgata	cactatcatt	acatttattt	gcagtattat	caataaatga	100500
atttgataaa	gggtgtaaag	gttccacgcc	aaataaagg	ataggcacia	actttgcctc	100560
atccgtgcga	ggtaaaaaatg	gcctgtcttc	gttgtatgat	tttgaaaagg	ttgttgcaag	100620
aacacaattt	cttgcatctt	ttaatgcctc	agctgatata	cgtagacgta	tactgtttta	100680
cttttgcctc	gttgtctggt	tattaaatcc	tccacacaa	gagtacatgg	cattatcgat	100740



attctccctc gcatgatcct ctatgaccaa accagccttg gtcattagaa aatgggtctgg 100800  
 gtctttaccg agaggaagaa tatctgcatt gaatttttgaa attatatcat tgacaacacg 100860  
 ttcacagtgc acagatctct ttatgaattt agaggctaga gatatgaaat gtctattttaa 100920  
 attggaacct agcctattcc aaggaagtga attttgatgt cttttattat ccctctcgtc 100980  
 agaagggaac aaaaattgaa gtaatctggt ccaattaacc gcagctgcag tttctctctg 101040  
 gacgaatggt tgacgttggt ttagtagctg aagagacact tgggttggt tgctactact 101100  
 agctggaaga ggggtgttgaa agccacagac gccaaatctg aacgtcatct ttagattggt 101160  
 gtgtagagat tttgtgtatt cctcttcgtg gatttttctt ctctttcat tgggcacaaa 101220  
 ctctttgtc ttttaattctt ccaagaagtc ttgtacattg tcctcattcg cgaacgttac 101280  
 ctctccatct tgccctcggt tcatcacgtc gttgttctta aacaggcgcg cagaaagctg 101340  
 catcagggac atttcttcca actttttagg atgtccttct tcctcttccc 101400  
 cgaagagata aaagacccta taacattctc aatgcaatga actgcatact gtctcaattt 101460  
 tctgtttgtc atccctgaac caaagcctat tactggaaca accatgccgc acttgtgggt 101520  
 gtatgtgagc ccacaagcaa ctgtcaattt cttcagaagg atctcgtatt tttcacttcc 101580  
 attaacatcg ttcgtaaaac tctgtttata cttgtgagca acccagttca ctataaataa 101640  
 caaggcataa gaagcagaca cctttcgtcc caaaagttta gaaatttgaa gggagagcca 101700  
 gtctgccatg tagccatctg attcgtgttc ggcgatagggt gtctcatca ttgtgacgga 101760  
 tgctttctcc actctatctg ctacaaattt cactgcttta acaatacttt ttctgcacgc 101820  
 atttttgtca ttagagttta ttcttgcata ttggagttca caggggtaat attcttcccc 101880  
 gtcattatct tcattggagg atgaattttt agttttgttg tcgttcttta tatactgttt 101940  
 caattttgag aagggagtat ccagtagctt ttctctacat gcactagccg cagaaagagt 102000  
 gaaaccgagc aggtaaagtt gtacacatga aaaggcacct ccaaaaatat taccaatgaa 102060  
 ataagcaaat attaatcccc tgatgtggtt gctagggaat ctaaagttgt gctccagact 102120  
 ataattcgtt tcaacgttaa agtacctttc caatctacga ggcactttaa tgtttatgtc 102180  
 ttggaaggatt tcaaaaattct cactaaacat acccgacaac ttgtttgatt ttcccaatac 102240  
 aatatctgac cttatgtatt tttcataaaa gtcataaacc gtcatcggtta aatcagttgg 102300  
 gggctttttt gcatttttaa gagtatccga aagatgcacc aaaccagttc tgtagtgtga 102360  
 gtttgccatt gtttcatcgg gcaataattt gctcacttca ttgtatgcat cgacaacttt 102420  
 ttctggagca attttttcag catataaacg ttgccatatt ttcttatcag gatctagtag 102480  
 ttcaacgtgt ctagattcta ttgtcttcac atcgacaaga aagtatcctt cacgtatact 102540  
 ggtgcaggcc aatctctcca tcttctgtt gtagtagtct atcctcttcc tgaagcaaga 102600  
 taattctgaa gcaactccat ccaacatgtt gtacagggct gattgggtac accactcgtc 102660  
 actatttact ccatcacgag cagtgaagaa atcttgaagg aaactgtata tagagatgaa 102720  
 cgaaaacagt cttctggatg atgttttcag tctcatgtca gttctctgca taagtttctc 102780  
 gtgccgtata actctgtcta cgaattccca tgtcagggtt tcattttacgt attcttttag 102840  
 gagtgccttt tctgttcta tagttgtagt agaagttgaa tccataatgt aggaatgagg 102900  
 gaggtacttt ggcacgcacg tgtgcaactg ccgagacaca acggtgtact gaacaaggg 102960  
 taaagacggt ggctgtttta cacctctcag aggaatccct ctaccagaat aagtcctttc 103020  
 atttttttaa ttaacacacg caaaactgac gctgaaatcg ttctgcaata atgggagagg 103080  
 aagtgggagg agaactttcg cctgtggtcg tagggatact ttccggtggtg ttaaccgcca 103140  
 taattatctc tactatattt gccatctact acttttattt ccgtccaaag gaagggttaa 103200  
 agaaatcgcg agtcagtaaa gacccttgg ctatacgcgt agatgctaac catgtcctct 103260  
 ccaataattc tccctacatg atgttcaata caaacattaa ggctagaaat acactaacag 103320  
 aatttgatcc caaaaatatg aatgaagagg aacgagcaga atttttcaag atccattatt 103380  
 ttactgggtt ccttaaagat gtggaaaagg aacaggtgca gaagaagaag gaggaagaag 103440  
 aagaagacga agaggaacag acgacagata ctaaccccaa aaataagaaa gtggaatatg 103500  
 taatgctcga atctctgaga ggggataagg ctggattttt gatagatgag aataagacca 103560  
 tgtctataca aggttctcct ggtacaaaaa ctgtctgtaa gaaaacgacc ctaccaaagt 103620  
 ttgtcccaac ccctaaaacg gaaaagtatg ttccttcaga tgaaagattt ttcaggatgg 103680  
 aagaaggaac aggagaaaag gcagaacatt attccttcac gtctgttaag acaggaagaa 103740  
 ccttaatcat gggctcttgt ggctaccctg tgatgaagga tataagcaaa ggtacaagcc 103800  
 aacagtatc gtacacattt gaaccaataa aggaaaaaaa gtaatgtctt tgtttatttc 103860  
 cccctttttt agaaaaattg gaacaacaac aatgtcgttg gctgtgacag aagattacgg 103920  
 gcacaatgaa aagttgatca aacggttaca aacctctgta tatcacaccc ctctattagg 103980  
 tgcagaccat gtaatgaaat ccatatcaga ctacataatt tctcgtcgat tcatgaacta 104040  
 cacaaattta ttaaaacaag ttgaatatgt ttctcgatgaa gaaacaggag cagttatagc 104100  
 taatatctgt ctgttaaaaa tcctagaaag atgcgcacag aaaggaggaa tatatgatgc 104160  
 accagaagat gttgcattct tcaattctaa gatgggggaa gtaacgcgcc tatttactat 104220  
 tataggagggt aggcccaata tgacgggtgcg gggttaattt aaacatgggc agacaaataa 104280  
 tctgcctat ggttatctca cagatgataa tgatactact actgttactc ctctgttac 104340  
 tctctctcca tctccagctg caagaagatc cccttttttc acacgcactc tcatatccga 104400

gtcgtcttca	gttgaccatt	atgtattgat	gcatgataac	ccaaaaagat	cttcatttaa	104460
ggtgtatgat	attcacgcag	aaacctttcc	ccataaagct	ccttctgttc	ctaccttccc	104520
ccctaaaacc	tcgtttgaaa	tttctgacgt	gactctcgat	tgttcaatgg	agattttttc	104580
accgagacagg	gatgttttag	acaatgttca	cgactatatt	gctaacgacc	ccgtaccatt	104640
tttagtggat	gttgtgcacc	gtggatctag	tctccgttga	gtaagtactt	tagttgcagc	104700
cggtctcatt	agacctgaaa	aatacccttc	atttgttaga	aaagtgtgct	tagatcaaaa	104760
tggcatccac	ttcttcatca	actaagaaaa	gagtacatga	agaagatgaa	aatctcatcc	104820
cacaacccaa	aaagaaaaaa	tcaaagaaaag	tactaccatt	tctgttgac	aagtatagag	104880
ctgtggataa	aaaggtggta	aatctcatatc	acaagatatt	agatcaagaa	aaggaccacc	104940
tttctagtag	cgaactgcaa	atgataactg	aatgtaattg	tgcgcgagaa	gatctgctta	105000
aacatcttct	agacgaagga	gaattttaacc	ctactataat	tgaagttagta	tcattccatgc	105060
ctattgaaac	aatatacgaa	atactctctt	cttctgctga	cgacaagaag	tttgtacaga	105120
tatcattatc	aatgttgatc	cacatacttt	tcttcgctga	taagggtact	atgtgggtat	105180
ccaacgcgtg	cgttcaaaat	gttttgggga	acgactataa	agtgggaattt	gaaaaataac	105240
gtaaaaagta	tctgatattg	gaagacttac	tgaacggcgt	ttcaaatcat	tggtctgaac	105300
atggctcctc	ttctcacatg	ctccattctt	caatccctat	tgtacaagac	atgttattga	105360
acaggctggg	gcgttacttt	agcacgtatg	atggagatgc	tcaattcgat	atatcattca	105420
taattaatag	tgtcttgtgg	ggaattgata	aaagtgttct	caacgaattg	acacaattga	105480
tatcgagggg	tgttttcatt	gtgtcgtacg	taccgatgcg	tgtacgaaca	ccttcaaagg	105540
acagtaatcg	gccacaaaat	actccttcac	aaaaatgtc	agcactaggt	atgaaactca	105600
atacattttc	atccagaatc	tcagtgtaca	gaaacaatac	ctttaaaaaa	ctaaccgagt	105660
tagtgcataa	ctttgattac	ggttccaaag	atgcatcatc	atcatctcct	cctcctcctt	105720
cattatcgga	cagcgtcaac	acttttgtga	ggttgtacac	caactatgac	atattcttaa	105780
aggtgatttc	cgactggaaa	atgccttatg	ggttctttaa	gaaaactttt	gacgtccttt	105840
attctaaggg	gttgatgaca	ttatcagtg	ctgaatatac	actcaaaaaa	gagttgggta	105900
cgtttttgcg	cgccttgaag	gaaagggaag	ttttaatcta	taaaatggag	aagagagaca	105960
ttatatgtat	actgaaaaag	tctttgttcg	gatttaattt	cagggtgtta	aaacaattac	106020
tcctctctct	caaacacttt	ttaaaaattg	aagagggtta	acatatagca	cgttttgtct	106080
ttagagatta	cagtctcatg	tgcaaaaactc	aaaaagattt	gcagagtttc	cctgccatac	106140
agtctgcttc	acttttcatg	gaagaattcc	cttggtctgc	aaaaacttgg	atcgacgacg	106200
atgatgatga	aggaggaaaag	ggacataccc	tattaacatt	tgctatagtg	cacagatatc	106260
ccttaataag	ccaacttatt	tcacacccaa	ttttaaaaatc	gttagtgaat	actacaatga	106320
gagacaagca	ctttactccc	ctcatgcacc	tcgccaacac	gtctataatg	taccaatgca	106380
atacactctt	gtgccttata	ataaatggag	ctaaaccaga	gttcataaac	aagttcaacg	106440
agaatgtttt	gcataatagcg	attgaaaatg	ttaactatgg	agtcatacact	gaattgagag	106500
gaacattatc	cagcgaacaa	attgaaaaaa	tggtcaatgt	aagaagaatg	atggataata	106560
caacactctt	aatgatcgcc	ttggcgaggg	agaattattgt	actcgctcag	ctttttgacg	106620
gtctttacaa	gccccaaaata	aagggtccgtt	tcggttcttc	aaagaggcta	aggataccag	106680
agtttgtcct	cttaaagggc	ctaaagggaat	cagttgcata	tttggaacg	aggaatatat	106740
cctacgatat	taacatcata	aaggatgcag	taatggacaa	cagtcttttt	gaagaggagt	106800
acgaaatagc	agcagcagga	ctgcgaggca	ataactgcga	ccctgaagca	gacgagaaga	106860
ctatgaacac	gtggaacttt	ttcaccaaaa	attcaaccaa	atgggcaagc	tctattttcc	106920
aaaagaatag	gcagaaaattt	gtaaagattg	tggatgggtat	gaataggaca	tatgaagact	106980
ctgaatgtgc	aatatgcttg	gatagtctgg	acggggatct	tccttcaggg	agaacaacgt	107040
gcggtcattg	cttcacacac	gtctgttggg	tatccttgat	aaggatgagc	gggccgaata	107100
atggcagccg	cgcaagagga	ggagggaataa	aatgcccgtc	ctgcagacaa	gtcacctgcc	107160
tcggaaaaag	actagggggt	gccgactatg	atattgaaac	agagggaagaa	cgtgacacga	107220
aaaatgtcgt	gccttcggta	gaagaaggaa	gaagggaatg	gaggaagatt	ggtgttgaca	107280
gatatgaatt	tcttgttaggt	ggagtgtgga	caaatgaaat	aaaactataa	aatgaactag	107340
aatattggta	ttatttttac	accttacaca	tgttccagaa	acgtctgttc	cagaaatgtg	107400
ttttgagatt	tctggacaag	ccattttctgg	agtaatctga	cattgggtcg	accagcgggt	107460
ccaccctcgg	aacttgacat	tgggtcgacc	gagcgggtcca	ccctctaaac	tcgagcgacc	107520
cagaaaaatt	tttaaaaaat	ttttgagatg	gagaattgagt	aaaatctctt	ccctgaaagg	107580
gagggtctgaa	gaggctatga	cggttctagg	tgcccgccct	ctccagaaac	gtctgttcca	107640
gaaatgtgtt	ttgagatttc	tggacactcc	atttctggag	taatctgaca	ttgggtcgac	107700
ccagcgggtcc	accctcggga	cttgacatca	ggccgaccca	gcgggtccacc	ctctaaactc	107760
gagtgaccca	gaaaaatttt	taaaaatttt	ttgagatgga	gaatgagtaa	aatctcttcc	107820
ctgaaaggga	gatctgaaga	ggctataaccg	gttctagggtg	cccaccctct	ccagaaacgt	107880
ctgttccaga	aacttactaa	agttaaagta	cgtttctgga	caagccattt	ctggagtaat	107940
ctgacattgg	gtcgaccacg	cgggtccaccc	tcggaacttg	acatcaggcc	gaccgagcgg	108000
tccaccctct	aaactcgagc	gacacagaaa	aatttttaaa	aaatttttga	gatgtagaac	108060

gagtaaaaca	ctctagtgga	ttagggtggt	aacagcctac	cctttctggt	cgcagctaga	108120
cccagaaaca	tctgttccag	aaacttaata	aagttaaagt	acgtttcttg	acaagccatt	108180
tctggagtaa	tctgacattg	ggtcgaccca	gcggtccacc	ctcggaactt	gacatcaggc	108240
gcacccagag	gtccaccctc	taaactcgag	cgacacagaa	aaatTTTTaa	aaaatTTTTg	108300
agatgtagaa	cgagtaaaac	tctctagtgg	attagggtgt	taacagccta	ccctttctgg	108360
gcgcagctat	accagaaaac	atctgttcca	gaaacttact	aaagttaaag	tacattttctg	108420
gagatgtgcg	cacgtatgca	tgcatgcata	cgtacctatc	tgtatttttt	taaaagtagg	108480
ggaaaatgct	atttaataaa	acaaccacca	ttctccgctc	actatacagc	tgaccttcgt	108540
tgagtaacca	tgtctagcgg	aaaagtaacc	tacgaaatcg	ttgaaggggg	attgttgaac	108600
aacaagtacc	ttctagatgg	agggtgcagca	atctgtctgc	agtctaattg	tgttgcaag	108660
aaacgtcacg	ccggttccct	ccacgataac	ctcttcaaga	tgctaggatt	tgcgaccccc	108720
tataaacaga	gacggggaaa	aacaaacagc	aaaaatctgg	ccataattga	agatagacct	108780
caactcgggt	cagtatcagt	tgtccaacac	ccgacagaac	cagaaagggt	ttgtcccatg	108840
acattcttat	ttgctcagta	caatatgggt	aatggaagaa	aatgttactt	ccctaaccgac	108900
aaagagtatg	ttgagagtgc	caagaagcac	gaaaggggtcc	acaaatcttc	cacagaaatg	108960
aaaagattgc	gcttgattta	ctttaacaag	tgtcttcacg	cgatcgccaa	atcacctgca	109020
atgaagaagt	acaacaagat	aatcttccct	gccagaattg	gggtgcgcgc	agctggagga	109080
gattgggaga	agtaccatgc	ttctatttga	gatttctcca	caatcattga	taaggagtg	109140
ataatagtgt	ctcaaaggat	gtaattaaaa	aataaaaccg	tcgctggcaa	ccccgcgcac	109200
caccaccatc	gaggcgggta	taaataaggg	gcgctggcac	atgggtggca	cactcgcact	109260
atgtcttcca	accgattcag	tcagctgagg	ggcaacgagg	agatggttgg	ggactattca	109320
agatggacaa	ctgtcaagaa	caggaggaac	agacagcaac	agtattccca	tagtttccgt	109380
ccccacaac	aacaacaaca	tcaaaaaaga	acatcaacca	attctcctcc	tgctccacct	109440
cctccattcc	ccatcattag	ttggggagcc	ctcggcagct	actcaatgta	tcgactggat	109500
gaccagtgc	gaaattgcga	tgaactggc	tattacaatt	tccactctta	tgatagaaag	109560
agggaaaagag	ttcgctcatt	aaacaacact	ccaagtgaag	gcatgtggcg	gcgcacaagt	109620
agatcttccc	ccttccctta	taagaagaag	gcagttgacg	aagctccacc	tcctcaatca	109680
aaccaacaca	tgtaccccc	caacaagtac	agtttccgtg	aatatactcc	ttcatcaaa	109740
cttgtgaatt	ggcgagaccc	ttcacaagaa	aaacaggaca	agatcttaca	agaggaagaa	109800
gctcgcgcgc	ctacacccac	tccccaaaga	aaggaaccag	aagtagaaac	taaagatgat	109860
ggtgtcatcg	aggaagaaac	tgaccagaa	ccagaaccag	aaccagcccc	agttccagac	109920
ccgatatttc	ccgcaataac	tgcaactact	actactacta	cagttgcaac	acgtcacgac	109980
gattcttctta	cagtatttct	cagaaatggt	attctgagta	tcgtgttttg	gtttctgggt	110040
gtttattctg	cattattttg	aaaatgtatt	agatctaaga	aggaataaat	aaaatgggat	110100
atgaaattta	aatcttttatt	gtttctttcc	aattactcca	ttgaaattgt	cctccttggt	110160
agtgcgcctt	ctggcaacac	tcccatgtcc	acaagtgtta	aaggttgtat	agttccatcc	110220
ctctttgggtg	ggaatgggtcc	tctatctgat	tgacaaacaa	aggggaatct	atccttcaaa	110280
tcaccctctg	aaacagaaaac	agttttttct	cttgtataat	aatatctgct	cattaatcct	110340
tcatatccaa	aatggtgatg	gtgtgcaaat	ttccttagat	taggtcccc	tacgttctta	110400
aagtagaggg	ttgaaagatc	atcaaatggt	tctgggcctt	taactcttac	ccctccaat	110460
gttggttttg	ttccactaag	atctgaagg	ctcatgaaat	ctgcataggt	tatttccaca	110520
cctcttgtag	agacactctg	agcagctgta	ggtgcgcgaag	attcgaatcc	tcactggtag	110580
ggtgtcaata	cagggaagtac	atggactcct	ctctgactc	tagggactct	tttactcgct	110640
tcatagaaac	cgaccatgtc	tttttgctct	ccctctacat	actttgatga	tgatgatgtt	110700
ctaacgcgga	gattttctgt	agcctttcct	tgcttattga	ttgtgaaaac	ttccgcatct	110760
tcttccaaat	ccacataatc	tgaatatagat	aatctcctct	tctttgtact	accggcagtg	110820
cttccctcta	acatttccat	attgctgtta	actacttcca	ccacaacttc	actcccatca	110880
ggtgggttcaa	tgcttgccgc	ttcattttatt	gtcgcagaat	cggccgatat	tggtggcgaa	110940
tctatgtctg	taaagcaagg	cgttgcatat	aggctcacta	tatccttgac	accctcaacg	111000
tagggggaaa	atgttccccc	gttaaatgta	gaggcagtaa	ctccatgaga	agactcgaaa	111060
aagtccacta	tccttctage	ctttgcaccc	tcttgacccg	tttttgctgc	atcccttata	111120
atgtggaaaa	tttcagcctc	cgcaggaggc	acgcccggga	agcgttccata	atcctccatt	111180
aaaatgtcat	atgtgtcata	atctgtcata	ccaacatccc	ctaacaattc	tgcttcccta	111240
tcgaacgtac	ccttcccttc	gtccaaaaca	gcagcaggta	ccttttcttg	tcttggtgaa	111300
taaggagtgt	aagccgtact	cctacgcatg	gcacaaaatt	tattaggggc	caacataaca	111360
ttggaagata	tatcccatat	attttctgtt	gcagctttta	ggtttatcct	tcttccatta	111420
gttttagttt	ctccgatagt	atttagtatt	ctagcgtctc	cggattgagc	tctcaggagt	111480
gctgcagaga	gatgaacatc	gttttttagg	agggtagtca	tgaggggaaa	acgtgaatgc	111540
ctatttccctc	tgctaggtgt	tgccgctata	gggtcagttt	tgttccctacc	ccctccattc	111600
actttgggac	agaataaacc	gttggtgcaa	ttacagttct	ccttttctcc	caattccccg	111660
attttatcga	atatgttttt	actattagga	tcttgcatct	caaaaatact	cttttctcca	111720

atattggcag	gaagtagccg	tgtttggtcc	tctacataac	tctttataga	aaatttagat	111780
aacgaagaat	gtgattgcgg	gcccaggggt	ttggggacca	tcattgtagt	gtactgagag	111840
agggcacccg	agcacgggta	tccgaaagat	tgccaggttt	cttccctgac	acaaggcggg	111900
gagtacacgc	cggtcaagta	gtggccggcc	gtgtgggggc	tcagacccca	atctgctctg	111960
aaccgattag	ggatcaattc	tggcctaata	aacgttttcag	aggaaagagt	cctaaccctg	112020
agtagcgact	caaagagagc	cataggagga	gggatgaaat	tgagggtccc	ttcagtttcc	112080
ttgacttcac	attgttgcca	taattcatcc	ccttcatgga	attttagagg	gactattcca	112140
tgcatggacc	cgaccgtgcc	gttcgagact	gattcttggg	cactcttttg	aatgagaaaa	112200
ctggaaggta	tgggtaaagg	aaatccccct	aaaactccca	aaaatgatgt	agcttgagca	112260
gccgcagaca	ctccccgagg	aggagcgaga	ggcttagaaa	agggtcgtgc	tggagcgtag	112320
gaattgaagg	agggtttaat	gtctgttgtc	gaatttagaa	tagctctaata	gtctcccctt	112380
gttactcgat	cataagtgtt	gttgatggct	gtcaagtgtg	cggccgaacc	aggggtctct	112440
ctcgcgttga	ccggcgatct	gcgttcgaac	caggtgtggt	cggggagaga	agatgcagga	112500
atttccatag	gtcttacaag	agggtgaagt	gcaccattca	ttaaaactcc	atctccagtg	112560
aggtattgctc	ctgcaaaaaa	atcgggcttt	tttgataacg	tttttgcgac	cctttcggcc	112620
tcgcgcatac	ccgtaccctc	tattccatta	tttagtctct	ttcccagaaa	aaatgtccct	112680
tcattagaag	ctatgcgcca	cgcagttgga	gggtgagcaa	acatttcata	tgcttctttt	112740
tcacccactt	tcttaaacac	tgcagatgaa	ccagccttgt	aatgatcaat	tattttggga	112800
gaattttag	acaatatatt	ctcactgaat	cttgcaagg	tggaactttt	cttagcacc	112860
tgcgctgtaa	caggttgtag	aataatacct	ccaatttctt	tgaatttggc	attgttttgt	112920
tcattgtata	ttgatcttgt	attgatgggc	aggccagacc	tctcttttcc	ttcgcgtgtc	112980
aagttgggga	aaaaatagtc	cagaagtata	ttgtaacgt	cgtcaaaact	attactagaa	113040
tttgtaacca	tcatagaatt	taatgtactg	ttcgcaaaag	attgggtctct	ttcttgtgcc	113100
gtcaaaagag	agtgttgtaa	aattgtctcg	atgggagtgt	gggccattcg	cttagcggta	113160
ggtctatagt	gtaccaaggg	caatatctta	tttactgttt	tatcattttt	agcagaagg	113220
ttttccttta	ttactctatg	gagtgaagga	acgacgtatg	ctgttgcgca	tatgccagag	113280
tcgtgagatt	ttctcagttt	ctgtaagggg	ttatctttgt	acgaatcgta	ccctgccttg	113340
tcagaatctt	taaggaagga	tgtttctatg	gtcgtatcgt	tgtgcgcac	ccttctgtca	113400
gtctttactc	ctgaaggatc	gcctgtctct	atgtgaaccg	tctcggggaa	gattttaccc	113460
tgggacgcta	ttatgcccac	agcatgatct	atgctatcga	ttcctaaacc	gggaatattg	113520
ctatggggta	aaggcacgac	cctgagcgat	gcgtttggat	gttcagtaaa	atttccccct	113580
ccagctccaa	ccagcccatt	atcgtttagt	gcactctcgt	tccagatgtg	agcggccaca	113640
cttctgttcc	cactgtaggg	cactacacct	cctacagatt	cactgactgt	gagttttcct	113700
tgttcgtcag	aatcttcaat	tcttgagtgc	caattgccgg	cggcgtccat	ggcagattcg	113760
cgcaataatc	ctccccgtaa	cagatggcct	ccgttaaacg	ctgtaatttc	aggtgctctt	113820
ccaaaaagag	aagcggcata	agaatagccg	agagccgaca	tcctcgtatc	gtttcctggt	113880
ctagggaggc	acgcgtcctt	agtactacca	cgtaaaagtt	cgtacctttc	tgttatttgc	113940
agcaataatt	ttgcaccctc	tacaaagggt	tgagcataat	ctgttttttg	cgcatcggga	114000
agatctatac	cggccacttc	ttccagtctg	atttgtgtac	tttcatcaca	aagacagccc	114060
tttgttccag	tatctcctac	aaccccatat	ctgaacatga	cagggcagcc	tagaccttgt	114120
tccttttctt	gtataaccac	ttctttcttc	tcacattttg	gggcgttgtt	ggcaggtgtc	114180
ctctgctggg	agtagccagg	ttcgcacaaa	caatccatag	tagaagtcac	actgcttatg	114240
tcgaaatcga	tatcgtccac	atcatcaata	attgagaatg	ttgtagggtg	tatagggtcg	114300
cccactcctt	ttcctccaca	cagtacaata	tctgtgcacc	caagaatggg	ctctaattct	114360
ggatctttat	tgagaaggga	aggatccttt	cgaccacag	acgaatctac	acaccctctt	114420
tttttggtga	tataagtctg	tccctttttt	cctcctccag	ttatgactcc	tccagattcg	114480
tagtgcaaat	aagcctcctc	gacagtagta	aaagggttga	cgtttctgta	ggctgcactg	114540
gtagcaacat	catcgtcgtg	cttgtcgaac	gcataaagg	ctgaaaaaat	agtatcggcg	114600
agagggacac	aaactccgtc	atattcatca	tcatcttcta	taccgtattc	agttttcaat	114660
tctctactgg	atttgacgca	agttcctccc	ccgagagaa	gacagcta	gcattcaaaa	114720
tcgttgccct	ctacgcgaca	tttttttaga	gttttaaaac	agtcaaatgt	gttataatgg	114780
gaactaat	ccccgggtaa	tagagattct	acaaaacctt	tccttatgat	gtctgacttg	114840
actccagcca	taccgaagg	ggcgctcgta	tcgacatata	aggggtcgaa	aggggtggga	114900
ataattcccc	cattttcttt	taatcctttg	acgcgtaaag	agtttcgaga	catggctcct	114960
gtatctacaa	aagagctcgc	ttaaccacag	ttgactcac	ccgggtacaa	taagacaaaa	115020
attataaatt	gggaagaccg	atacagtctt	tcatgacaaa	ataccccgag	aacaaaagat	115080
gtttgtctag	gaacaaagaa	acattaaaaa	tggttgcccg	aagctccaag	accaaataccc	115140
gccgtggaag	caagaagagg	tccaccactg	ctggagcgat	ctccaagcgg	aggagcccat	115200
caatgaagaa	gcgtgcagga	aagaagagct	ccactgtccg	tcgcccgttc	tcaaagagcg	115260
gaaagaagtc	tggagcccg	aagtcaaggc	gttaattctt	ccctgtacaa	caactatgtt	115320
atthaattga	ttttttttct	tctgaataat	tggaaataat	aaaacatcca	ttgaacttta	115380

tgcagatttt	ttattcaatt	ttttaaccaa	ctataaatcc	acatgtggta	taaaagttta	115440
agggtagaca	tattattata	gatgatgtaa	cagatgcacg	gtcaataaca	acagggcctg	115500
aagaagtgtc	tatgaatttc	ttgtagaatt	cttccctgtc	tgtacttgac	gctacattga	115560
caaatccagg	agggctact	agtaacacct	agctagtgc	gggttctatt	ttagacatga	115620
cattttgagg	ctttttaaca	cataatacaa	tatcatcatg	gtcgaatgtc	gtcctgttcc	115680
agagacggcg	atgtagtgtc	gaattttctat	ccactaaaac	tagaccatcg	tcatttttca	115740
aatgtcttct	tccttcattt	ctagctacaa	ttttcatgta	gaatcttact	ctttctgcat	115800
acgacaccac	cttgtcttga	ttcccaacta	tagcaaatag	taacgaggaa	gtaaattcat	115860
tgtttgcacc	attattattg	tgaacccgt	ttgcactacc	aattctctca	catctagggt	115920
ctagacgtac	cagaaattgt	ccatcggcta	accgagtaac	agggctgcgc	ctaggcctag	115980
tggctagccg	agtctgtaga	ggggcaactg	gttgtctcaa	ttgagaataa	gaaccactg	116040
cgcgtaggcg	gccagtgttt	ctccagttcg	tgtatctgtc	tgcaggttcg	tcagggtatt	116100
tatttgcctc	aaaaacaaca	ataccctatt	ttcctaaaat	agccgcaagt	ttcttgtctt	116160
ctttattttc	cagaagattt	tcttttttgg	ttccagactc	gtctaccacg	tatgtagtat	116220
ttgaagtgtg	cacaatttgg	tctaattgtt	cttcagaaac	gccgtctccc	atgtacattt	116280
tagcgtacac	tttctcccc	tgtgtatacg	cttcactagg	tgtctggcaat	aactgtcgat	116340
gacgttcttg	ttccttatcg	atttgttccc	tcaaccagtt	ttctattgaa	aattcagttt	116400
gagtattctg	attgacaatg	tttgtagtat	cattcgttct	tatggaagaa	atgaaaatgt	116460
gccgagggta	ctttatttct	ccccctaat	cagcgaatgc	ccattcaggt	atgcttgtag	116520
taggggggac	aagactagtc	tgttttatgc	aagggctccct	tatccctcta	ttagaatatg	116580
gaggaagtgt	agtgcgcacg	tagactatgc	ccttgttgtt	tatgtcgggtg	gcgggtgtgt	116640
cccttacagg	tgccgtaaag	gttgagggat	cattccatat	agccaattcg	ttagatggta	116700
cagacatgtc	tgaaaactgt	ccttcaaaat	gcccgcccat	cgtaaagggc	atctgtatgc	116760
ctttatctgc	acctccaaat	acttccatct	ttacactgaa	aacctatacg	tctctgtcac	116820
cgtaataaaa	tatgggctcg	cggtttatct	tgaagaaag	aacgttttct	ttgcaagta	116880
tagaaccaga	cacgggcggg	gagactgtga	tgtggtagct	tttatcttca	aggcaacat	116940
catcatcatc	atcattgacg	ataacatcag	atacgggtcaa	gacagtagtt	ttgttcttgg	117000
acgtgatata	gtgaatgggt	tggaaatgggt	tatcactcgc	tttactagg	ccgttatgcg	117060
aatctaaaac	agggaaaaca	gcgtacagtt	tagtaggttc	tgcgggtgcc	ttaaatgtag	117120
gctgttccac	tcccagtaat	gcagtttgaa	cggggttaat	aatatctcca	ctctgtactt	117180
cagtcaatgg	accaactgaa	acgtttaaag	cacgaggcag	tagagatgtt	gtgcgttcac	117240
cggcataatc	tgtggcgggg	gaagaaagca	cagccgttat	tcgaggttct	ctattatcat	117300
tatcagtttt	caagatgata	catggtctta	gttggttaga	gaagaaatct	ggcccgctcg	117360
tgggtccagg	ccaattcggg	gcagttgcag	catctactct	aaagtttacg	ctagaaccga	117420
agcaccgcag	tctcgaaggt	gtattatgta	caacaacttt	gttcgtgtta	gcattcacat	117480
cagtaacttat	aaccgagaac	aatttatcag	attgagtaac	ttttgatgct	ggaaatacaa	117540
tattgtaccc	agaagatgtg	gttaaaacgt	cagacgacgt	tttgaagccc	acaatccagc	117600
ccaacatttt	ccattgcgct	tctatttcac	tttcaggtag	agacacacat	ataaagggtc	117660
gagctggggt	agatggattt	ctaagggtgg	ttatctctgt	ggcgtttaga	atcttgttaac	117720
cctcggattc	atttccctcc	atattcttta	acgthttact	gtatagtact	ttgtttaata	117780
ctggtgtggt	aacagtggtt	gttttcaggt	acacgaacga	cactgaaggt	ttgttaatat	117840
ctgaaagtgt	atttttcgca	ggtagttag	ctgctaaatt	ggacgctatt	tcagcttgct	117900
gggaaagacg	gggaaaatta	gcagggtaat	ctgaaacata	tttttcattc	aatgtatcag	117960
gagtgtatcc	cacattatca	tagtttattg	cttttaggcat	atcaagcctc	atgagggcta	118020
catttctaca	agcaaaaact	tccgttattt	ttggtagtgg	tgaagatat	gttacctgag	118080
aatcgaataa	ccaggtcttg	ttttttctct	ctgttacaat	ggaaattttc	agctctatcc	118140
cagtgttgaa	aatacttccc	tttgttgata	gtgttatacc	gtcagctaga	gaatttactt	118200
ccacgtgttt	gaatacggct	tgttgtttat	tgccattggt	gcgagagaa	gcattcgggt	118260
catttgtgtat	aggaggccaa	cctgctggga	caaatttcagg	gaaagtttcg	aatagagacg	118320
gaccactatc	gtcgtctcca	ccccctctt	ctccttcata	tttttcatcc	actaaataca	118380
cttttatctc	tgtccttaaag	ggaagaggga	cattatctgg	taatatctcg	tcaagaatat	118440
gtatcctcac	gatggtatct	tttccacctt	tttctaataa	ttcttctgta	cggatattaa	118500
tcattcgaat	tggcagcaat	gatgatgatg	atgaaccatc	tctatcagga	gccgattgag	118560
gagaaatatc	atcaccaata	tcataaccga	tctgcctgac	agtgcgtttc	attatactgt	118620
gtaacaattt	tggactcctgc	acatcagtag	cacgthtttaa	ttctctcag	gatatctctt	118680
caatattttt	cgcgagtttt	tttgacgcg	gttctaacc	tttcttcaag	ttcttttctt	118740
ttttcgggtc	ttttctaac					

aaaaagttac	ccgggcatgg	cagggaatag	aaccaggttc	gtatcgtccc	tgattgctaa	119100
atgtatatca	gacgtggaac	aaggcatgga	gtgtgtggtg	agacaagcac	aggatgcatt	119160
aatgactcgc	ctagccaact	taaaattggg	cgattctctt	aaagaaactg	atgttaattt	119220
ggaatacttg	agatacgcgt	ctacgcccc	ccttggggaa	ttaaactacg	acaaacaaca	119280
atatgcgcca	acagttgaca	tcaacctaat	ggctcatttc	tcctacgctg	ctttgggtat	119340
agaaagtata	ctgaattcta	tacggagagt	tgtagtggct	aatcatcaac	gtagaaataa	119400
tggaaaaaaa	ccttctgaac	caatctcacg	ccctcacccg	ctgggagggg	tagaacctcc	119460
tctatcgtca	gagttggcaa	atgcaataag	ggacaagttc	atcagcatgg	gggcgttgga	119520
cagattgaat	tcagcaatag	tgacagcggc	cctgggggct	attgccagt	aacgtgaact	119580
attcttacgt	gaaaatgctg	taaaactacat	gtacgatgta	gaatttgag	aaagagatgc	119640
tgctactaca	gatacaggga	atgtagtcta	tccttccacc	aaaatggacg	aagatgaaga	119700
tgacataata	aagcgttcag	aaatattaga	taaggatatca	aaacgaccgc	caaaggaagg	119760
tatagactgg	cgccccaccc	ctgacaattc	gttcccttac	caattgattt	ggggcgatga	119820
ttctgtagat	gatactgttc	ttatagatct	catcaccaat	gcgatcgtgc	ctaataattt	119880
tatggcaaaa	tttatcctgt	tcatatgtaa	ccatttaagg	gcagttatta	ggagtagtag	119940
ggaaatttta	tacggtaaca	tttcttcttc	atccgataat	tattttgagg	atggacgtaa	120000
atggtgcttc	tggttgaacc	tgtacaatag	actggaatgg	ttcatgttag	tagtttagatt	120060
tgtaattttt	ctccactcaa	aaaaggagtc	cttttcagga	gctgacaatg	ttaacgtgaa	120120
aagacttctg	gtggtagttg	tggagagttt	tcctcccgtt	ctcttggaca	ctgaatgggt	120180
caagactaat	ataacgtcat	ggcctgttat	taataacagc	aataataata	gtacactccc	120240
tgtagacaga	gacaccttaa	tgagactagc	gataaggacg	agtagcgtg	cccgacatcc	120300
tattttcgac	gaaatttaact	ccttgacaac	agcagtgacc	aaccgtatta	ccttccagtc	120360
tgacagaatt	tgacacaaaga	ttttgctcgg	gcgagctctg	gacgaagaag	aagctggaac	120420
aaaaatgcta	gtaaaatcag	tcaaagagac	gggagaagaa	aaggataaga	acaatacgtt	120480
ctcttctatt	ggtttattac	tgaagaacac	aaaaaatgaa	gaattggaaa	taaacatagg	120540
cgataacgat	gatgagacta	cagatgtggc	ttgttgggca	cgtacttcct	cgacatcctt	120600
tatccgtaat	aggacatatg	cgtttaaaaa	aatatggggc	cttgaggatg	caagtgatgt	120660
agtcgagctg	aagcgagaga	gtgacgccat	tacatccttt	gtcaccgata	agagcagtc	120720
tctcctattt	ccgtatgtgt	ccgactggag	ttgcttacta	ttacatccct	gttgtaaagc	120780
accggccata	attaaaagtg	tgtggttaca	aatcctgaaa	gatttttccc	aggaaaatat	120840
aaaaactata	aataaaaagg	tacaatctct	ttcatctgag	atttgcaga	aatcaaacga	120900
ccgttttaaa	aataaaaaaa	ttgtgcccga	acacgttcgc	agtgtaaaaa	agttattaaa	120960
tacgataagc	aacagggagc	aagaagcagc	actatctaca	gaacactgta	tttggttaac	121020
gattttgtgg	aaacaagtcg	ttcagaacac	tctcaacctt	ctggagaatt	ttcccgtata	121080
aaaggggaaa	gtactcaaca	tgttcactca	cttgaccgca	gccttccgca	agatgaacaa	121140
cctagtcagt	cgagcttca	ttgacgttca	ccgagtggtg	gccgagttat	cctatccaga	121200
attcgaaag	gatgtaaaga	atccggaatc	atccatatat	agaactccga	tatccctctt	121260
ccaaaacaag	gatattgtta	caatagttgg	tgattacatc	ctctctccga	agacggactc	121320
attccaagtt	ctatacccaa	tcaagaaggt	catcgaacac	ttcccagtaa	tcttccactg	121380
caccacacat	aatgccccct	tctgggtaca	ccttctggac	gaacgccatc	atcgccctgt	121440
ccagagcctg	ctgacgtacg	agattgtgaa	tgccaagtac	aggggtattg	ttgtcatccc	121500
atactacagg	cgccccatca	actatcaaac	tggggaagag	ctactgatga	gcaaactggc	121560
gtccgtaaaa	gttctggaca	ttttaatgag	atgtggatca	tacaaattca	tctcattaat	121620
gtgcatgatc	aacaagaaga	acaacaccaa	ctttcttcac	tgctgtgcaa	gtaaatgggg	121680
agaagttgga	agcaagatga	tgctccacat	tgctgaaatg	ttctttgcca	accctactac	121740
tagccaacac	ctatccgacg	ctagtagttt	ccctgatgct	gcagcagagg	acgacaaggg	121800
gaaaacacct	gcccacttag	caatccaaga	agataatgct	gatgcactcc	tgttcctgat	121860
ctccctctac	ggcgcaccc	ggtttcaaga	taacaactcg	tacatgaaat	ctgcccttga	121920
actcaagtct	aacaagtgtg	tcaaggtact	atcctttgca	gctgacaagt	acgagatttt	121980
acccaacatt	aacaacaatc	aactagaacc	agataccatg	tgtggagtgt	gtgcaacatc	122040
tgtggaagaa	gatgaaaatg	aagggaaaac	aacaagtctt	tcctggtacc	agatgaattg	122100
caagcattac	atccattggc	aatgcctgat	gggaatgtgt	gctgctgctg	gcaatgtaca	122160
atgccccatg	tgccgtgagg	atgtggcgca	cgaagtactg	gaaagatgcc	ctcctacaat	122220
atttagatgg	ttaaaactgg	ctgagagatc	tgaacacaa	cgtgtacttt	ttgaagcaaa	122280
aaagcaagaa	ttctataagc	agatggaagc	aatgaaacct	cccagagttg	ttgttcctcc	122340
tcgcaggaca	tttctcacc	cagccagaag	aggcgaacga	gccatcagaa	tcgcaagaga	122400
aattgccacc	aacgccatcg	ctgaagccac	agctcaagga	gatgtcaact	cctacttccc	122460
tggttctcatt	gacgggagcg	gagaagaata	tgaagaagag	ggagaagaat	tcttcaattc	122520
tgaagaggag	gcgcttgctt	ttggaagacc	atcttctggaa	gatgaggaag	aagccagaca	122580
aatacagatg	cgccagtttg	ctgaactgtc	tagacgaggc	gtttctgtca	atattatta	122640
caatgataat	cctcatcgac	acatctctac	agtaaataat	gtgcaaccag	tttatggagt	122700

tgaaaagtca	cctgctgctt	ccttcaceta	caacatgctc	aagaatgacg	tctttgagtc	122760
tatacgctca	agagatactc	gagttggagg	agaaagagtg	cccgtcatga	acctgtccaa	122820
tgacaagag	gcattattcc	acgcagcttc	ttccatgctt	tgtgactttg	ccacagaaac	122880
aaactctcaa	attgttggat	tggactttca	agcagtgtat	gatccccatc	acatatccaa	122940
ctatatcgag	acgtttggta	gtcctcttca	cgcctaccca	ggagccgtca	cttttctgga	123000
cggggcccg	gactattatg	cagagagtat	cagatacgac	aatgataattg	tctcattctc	123060
agaaatggca	agtgaagtgc	acatcaccga	agcattagat	gtctttgagg	gtagttttatt	123120
atccccactg	ttcaagaaaa	tcaggactgg	aaaatcttac	tctaactgga	acgaccattt	123180
gaggcgtaga	aattatgctc	gagatattgc	tgaggaattt	gttaggggtat	gtgaaaactc	123240
tctagcttca	cgcgaacacc	cccctgttca	tgtacatccc	tttagagatg	gagcaatccc	123300
cattctcatt	gaatatatag	tagatttcat	ccaccactgc	atcacctggt	ctatgcaagt	123360
taatgcactc	cattgtatga	gaaagtacat	tgaacacgag	aatacaaatg	tgcacctgtt	123420
aaacttgctg	cctactgatg	aaaggggtga	agtttttaagg	gtatctcaac	tcagatggag	123480
ccgcttgctt	aatgaacaat	acaacactag	aatgtccctc	agcaccaaaa	gattgagcct	123540
catgaagatc	ttcaaccatg	atttgggtgt	gtctaaattt	ggtgtataca	aactcctaga	123600
tattattgaa	atgtactggt	ttactttaat	ctaaacaata	aaaataatgt	aaagaattcc	123660
atgtttttca	tccctcctcc	tctgtggtgt	gtatatataa	ggcgggtagc	ctcctccaca	123720
ggcacacttg	tccagtagcc	gttgctcgagt	gaacatcatg	gcagcagacc	tcttagagtt	123780
ggctatccag	gaaacaatcc	agtctgaatt	ggaagaaatt	gccgatactg	aattcctcaa	123840
ttatcttccc	cataaaactg	gcactctgca	agaagctgca	gctaattggac	ggccatatct	123900
tctacacta	gaaatgagga	acgaagttga	ccatttctgg	tccaagata	acaggaagct	123960
gaaactcttg	gggcatTTTT	gtggcaactt	gtatgtggag	gcatttatag	ctggttctat	124020
agatgctgaa	acgtgctggg	gctttttgag	gtcgcaagca	actggactag	gataccctct	124080
attgaaaaaa	ctagccctga	ttgcccggtg	ggataaatca	aatacaacta	attacaactt	124140
gtacattgat	agaaattcaa	tgatgaaaca	agtttttagt	gctgaaattg	ataagcgtcc	124200
ttcatcaata	cagaacactt	ctcacacaaa	atcttccccc	gtgtacttga	aactgatcga	124260
cagaagaacc	gagtgtcttg	ccctggattg	gtggacgca	tccaagagga	cggccaagga	124320
aattggagca	gccagaaagg	tttgtttcc	ccagaacctg	attgttgcca	tattaatacc	124380
tgcatacaca	gaaacgtttg	ttcttgatac	aggcaacgaa	ctagaacagc	aagtattgga	124440
tgatgcatat	tttaattg	agaataaaga	taaagtggat	gaaatgtgtg	tagtggccat	124500
attgagtact	ttgcacaatt	tatttgttag	gaaaagtctt	ccccatcatt	tgtacaatgc	124560
acctttccgt	cttctccct	ttggacaaca	ccatttcatc	aacattgaaa	attctctatt	124620
ctttaatgaa	gatacgacac	ctatttctagc	gtcaatttct	ataccatcaa	gtatggtaat	124680
aaaacaccat	acgcgaaaaa	atagtagatg	gagatgtccc	aataacttga	tgactgcagc	124740
agaacggtct	atatttttac	gtggagtttt	aacagtcagt	ggagattacg	gatgggtttc	124800
tgtaattgta	ggacttacta	taatgcctag	tgtacttttc	tatggcgacc	gaaaacatct	124860
aataaatata	gtcaaatcta	ataacttttc	tgccataacc	tgttcttatt	ggaacaagta	124920
tatggattgc	agatcttatg	ggtttgagat	tatagacaca	cctgaaaata	actgtgggtt	124980
tcgtataagg	gctgcaattg	attgctcgaa	cacagatttc	cattcaccgg	taacgcgagt	125040
caacaagaag	aaaacgagca	ttattaatgc	ggtaaagaac	ccttttttta	ttagacacac	125100
agaacctaa	tggtacaaca	aaaatgccat	gtgtgggtgaa	gtattggaaa	atgttggcgt	125160
gaccttcgaa	caacacgtcc	gtgttagtga	tgagtatatg	gacagatttg	gtagcttatt	125220
acttgagca	gaaaagaaat	ggacgtgtaa	ttatctagat	agaataaagt	ctctagaaac	125280
tatttctaac	aatctcaagg	gtaaaattga	caccatgtgt	aaaattctag	aaaccaagta	125340
taactacaaa	tcttcttcc	tatactataa	gcaaataact	gctacgtctg	atgatcctat	125400
aaagatgaag	attatcgctt	ctataaacia	gaggagggtat	ttgtgtaata	ttctagaatt	125460
tgctataatt	tcattcagaga	aaaaggatga	ggtggaggaa	gatcacacta	aaaccggcaa	125520
tggtggttgt	gcgtttttcaa	agtataagaa	aaaacaactc	gaacctaaac	aacatttaac	125580
tgttaaagt	aataaatata	ttgaggcttt	ttcgtttaatt	aagatgctga	gaaatgattg	125640
cgaacgcaac	aagtgtagg	ttaaagaggc	tgaaattaga	gagtgcgcca	acgaactggt	125700
acgagaactg	tatagggtct	cggccagaag	ttatgtccac	gatctggtat	tgaagcgaac	125760
taatgtccac	ttgacatggc	aacgccccta	cgatgaaaaa	gctaacacta	tcatgtcttt	125820
aatacctaaa	tgtaagctac	acacagtatt	gtacgataaa	gattcgcgcg	atgttaagtt	125880
gttgaatttt	ctgagaacga	gggatggaaa	ttataaccca	ataagacact	ctatgctgga	125940
gttagtatatac	ggagaggagt	acgctaaaga	tgtcagtact	gttacctgtt	ttgaatggtt	126000
aaaatggtgc	tcgaaaaaag	gtgtgattaa	atatgaagac	tttttggaac	gttacgagaa	126060
aacgggggag	gaagataaaa	acgaaaggga	attctttaga	ctaaaaaaat	gtagtagaga	126120
tcacactaag	gatataaaaa	aaatagaaaa	tgtactaaat	tctgatacac	tttattctta	126180
ttctctcgat	aaaaatgtgc	aaaccacgc	atcttctagt	acagttgtaa	aaaatgacac	126240
tgacggaaaa	acgtctatgg	tgggctggga	ttatatTTTT	tcaatcggtg	aaggagaaaa	126300
aacaaccaa	aaacgaaaac	tggaaacgat	agatatatcg	agtagtgacg	acgacgatga	126360

agaagaagag	gaagaagatg	aaggaaaaag	aatgaaaatg	aataactgca	gcagcagcat	126420
caagaacaag	agcaagaaca	agaatgggag	aatgtgttgc	acagatattc	tcaatgttgt	126480
agaaccttct	ctaccttaata	ctttatcggt	caatttgtta	aaaagtatgg	atgtgttgaa	126540
tttgttatga	tgttaaaaat	aataaaaaata	ttgtacattt	tattacattt	gtttgttttg	126600
acacctgagc	ggtttgggcc	tgttgacgtg	cctaactttt	gtcacccctc	atgaatacaa	126660
tttgtaaaag	tgctgaaatg	tacttgtttt	ttatccaaat	ttcctgtact	gaagaatatt	126720
gaaagaagac	ttcttgaaga	ggaccgataa	aaaaatggcc	accttcacga	ctgacgccga	126780
tttcttgctg	gtgggggatg	atactagtag	atatgaagaa	gtgatgaaga	cttttgatac	126840
tgttgaggca	gtcaggaaga	gtgatctaga	tgaccgtgtt	tacatgggtg	gcctaaagca	126900
gggatctact	tttgcctcca	atggaggcat	cgaagaattg	cgtcttttga	ctggagattc	126960
aacgctggag	attcaacccta	tgatttgtgcc	aacaacagaa	taaaaataag	acggtgacgg	127020
gagactaata	tctttcttag	tttcccgta	cggtgaaaat	gttggttatt	tcttccttat	127080
gtttaaaaat	ttgtcttggt	taaaaaaata	aaacgaaaac	tgtcaatata	ttgttttatt	127140
gatatacaat	atcccttttt	acacagaaat	ggcatgtctc	tttacacgct	gttgttcttg	127200
agtgttgttt	ccatcttctt	cttcttctac	gtcgtatca	aattcattgt	cgttttccca	127260
ggctcttttt	actccccttt	gtcctcggtc	actggttagac	tggtctccat	tttctcgctt	127320
cttttggttc	tccctcttct	ttacctcttc	aaaccacttt	tcggtcttgt	ttctcgtaacg	127380
agcttccttt	gtgggtcaaat	gtctttcctt	gttgtagcta	gtgacaattt	ttgatgttgt	127440
ctccatcatg	gaatggagct	tcctttgttg	ttcaatttca	aattgggtca	tattttcttc	127500
atcaagttct	tgaggcgaag	agatgtctat	cctctgccca	tcttgataga	cggtgccatc	127560
attcataaag	attcggatat	cagaacctgt	cactttttgt	acaataacct	tcaaatcttc	127620
atcacttgga	ggggtctgga	aaactatact	tgatcgggta	cattttaata	ccttgaactt	127680
gttttggtgt	ttgatactgt	cgttgcaagc	gataggtgcg	tcacagccag	tcacggccat	127740
gatgagtctg	ggcctcatat	ctgcagtatc	ccttttcttg	tttatggggg	agaaaatgta	127800
cctcaaacca	ttttctgttg	cagtttcaaa	ggaaccactt	ttgtctgtcc	tgcaacaaca	127860
gtagaattct	gtagaggata	tcgagaagac	gagacaagag	ccgctgtcgt	atctgaaggt	127920
cgcattggaa	gatttcttgt	gtacatttct	gagagattct	ttttgttctt	cagagacttt	127980
gaaaagtgtg	tcagtttctga	tatacgagtt	tttgtagctc	tctttcttta	ctacttgtgt	128040
ttctagtctc	ttttcaactt	ttgtgggtct	tactacttgt	gtttctagtt	ccttttcaac	128100
ttttgtggtc	ttttcgacct	ggttttctag	ttcctctttt	atttgtaaat	ttttcctagc	128160
caactcttcg	cgtatctggc	gttgcttttc	atcactgtat	aagataacag	cgtaggctag	128220
tgctgaaaag	tacatttcaa	tgcttgcagc	agaagggtgt	gcgctccttg	caaaaacgcc	128280
cgcaccata	aacaagggaa	caatcaagca	atgtgaagaa	tggtcttga	aaataggatt	128340
ggctatatat	tctcctttgt	aatgggtacg	gtgacaaatt	tcaaccaatg	ggaccatcat	128400
atccttgaaa	ccatgtttct	gcatagttct	ctccattgac	tttaaagcgt	gcgtgacgtt	128460
cattgtactg	aaacctaaac	cgattaagaa	atcggaactt	attatcatac	aatcgtgttt	128520
ttttctact	gcttcagaaa	gactattgtc	catgtacctg	tataattgac	tgaattttaa	128580
gtccttgtct	tttcgtttta	atttttctct	gactattgta	ttgattttac	gcacgtcttc	128640
ttcttgcggt	gcaagccatt	cctcgtcagt	aacgtgccca	gcatcatctt	tgaataaatg	128700
tccaggttca	ggaggggcac	aacttatccg	ttcaattctt	aaagagtcac	cttctacca	128760
ttcttgaat	ctttttaacc	ctcttgogaa	tttgagttgt	ttatctttct	tcacatgtac	128820
aacaaatgtg	tcttcggaca	attcgctttg	tctttatagat	tcgatagctt	cataaacacc	128880
catgacttgc	tcataattac	ttgttttctt	cccgatcatc	aggaatgacg	atcttagttt	128940
taactccttt	tctgatacac	tttcattact	aacagggtta	ttgtttttct	ctccaatgta	129000
ttccatattt	atcttcttca	aaacaacaac	ggaaaataca	gttttctttt	aagaaatata	129060
ggttacaata	ctgtggattg	ataattttac	tcacagcttt	gttcaaaact	taactacgtg	129120
tttcggccca	ttttgcacca	gatctgtttt	aaaggaactg	ggtaaaatac	acatgaaaac	129180
actccttttg	aacaaattca	gatttatatg	tatttgtcga	tttttattaa	tatttaataa	129240
gtatattttt	aaaaattagg	atggggacag	atggtgtaga	ttgatacagt	gaccgtccct	129300
gttatttttac	cttaattttt	tttccctata	atacaataaa	ttggtatcag	tttaggaaaa	129360
catattttta	aaatgaaggt	tatcatatca	actaaacact	tttattttacc	cccttttcac	129420
ctcagaata	tcttcacata	atcgcatgca	agacttgcgc	atttcttccc	tgccgtacaa	129480
atcaatagca	aacgctaacg	ccaacagata	cgtctccact	tctagagatg	ggggttgaac	129540
attacgggaa	aacatgtacc	ctagatacac	cagtggaaata	actaggcaat	gggaatgata	129600
actcttgaaa	atgggggtta	caacaatata	acccttgtat	tttttgagat	ggcacacttc	129660
ggccacaggc	accatgtat	atttcagacc	acttttccgc	atatttccct	ctatcgctat	129720
taatttttcc	atgacatcat	cagcactaaa	accgagacca	atcaaaaaat	ctgtgctaac	129780
aacaaaacgt	tcttttccc	cgaaaaggtat	ttgaggaaga	ctttttgtta	tataactgta	129840
agttggactg	aaataaaacc	ctctaaacaa	catccgtgtc	ttttcttta	cgagcttatt	129900
tacttctctg	acgtctttgt	cttgagaagc	aagccactcg	tcatcagtag	ttgtagacgt	129960
gtttttaagt	ataacgtgtc	ctttctcagg	cttaatagag	gataataatg	attctatctt	130020



gaggggaattg	tccccctgtga	gctctctttaa	ttcttttcagt	cccccttcaa	aagtaatgtg	130080
tttatttttcc	tttaattgta	caatgaaaac	gccgtctcta	aattcactct	ttcggatagt	130140
ttcaactgct	tcaaatgtac	tcattacttc	ttcatacttt	tcatattgtt	tttctccaat	130200
cattagaaaa	gaagagcgta	gctgtaactc	gtcttcttca	gggtgtctcat	ccaggagatt	130260
tatcagtttt	tgttccccta	tatactccat	gggttatttta	gtcagttgat	ttatttttct	130320
ctttgacaca	gtttttatgg	tactctggga	ttactttact	caaaaattta	ctcacatgtt	130380
gttaaaagtt	tgagtgtatt	cttttgtggt	aaactaacia	tgaagtacat	ttcttagtaa	130440
atgaggtagg	tactgattcg	tgatggataa	cgatggcgtc	ccatgtctag	tgatgtgagt	130500
cacataataa	aaagggaat	aaaaataaga	actagtaaca	tattttttat	tttatccttc	130560
tcaatatgtt	gtttgatttc	tttctaiaaaa	tggctgtctg	aaaaatggac	gctatccttg	130620
cggatattaa	tggaatgat	acagatttat	ccaagctaat	cacagacgtg	attcaaaaaga	130680
gagccaaggc	tgatcatggat	agaaataggg	ctaaaatgga	catgaataga	agagtagatg	130740
aggctattca	ggaagccgta	goggccaaga	aacaaaaagc	attagtggta	tttgataaac	130800
tcgtggaaga	aactgacagc	ggacaaaagt	tccttccaac	attatcggga	tcggattacg	130860
acgcgtgggt	agacagagcc	atgccttcgc	atatgtgaact	tgtagagagt	gttgagggag	130920
attcctttgta	tgataaaact	cctcctttta	acgtacaaga	catagacgac	caaatcggtg	130980
atgagataga	tacaccaata	tcttaccttg	ccatggtagt	ggtaaaagtc	gactgtgaaa	131040
ctggggatat	cgaagaagag	tacaatcttg	ctctacctt	tggtgtgaca	caaaaataata	131100
aaatatacag	agatgaaaga	gaccagattt	ttacaaaggc	tgataaatct	gtgcgtattt	131160
ttaaacttgc	taaattggat	agtatatcag	gtaaaagtag	acaactgacg	tatgcggtaa	131220
aaaataacaa	tgaatatata	gagtttgtct	gtagcgtctt	tgcaaggttt	gaatctgaca	131280
gtgacactac	taaatcaggt	ataggtattc	gtgaatatga	caaacctaaa	aatgaattcg	131340
aatatgagga	gcgagagatc	tttacatttt	ttattccaat	acaaccgcga	gggacgaaat	131400
tattgtttata	ctttttggtg	gacgtcaggt	cccgtattat	ataggactat	aaggcacaac	131460
actctgtttg	aagagaaaaa	atgttcgtca	taagcatagc	tacgtcattg	gtgttattct	131520
tcttccttct	tttcgtctca	ataacaattt	tagtgggtgc	aaaaacaatc	gactctcaac	131580
cttttaggaa	aaggagggaag	aggaagagat	atcgaccag	tgagagtgg	gacggcatag	131640
acggaggaac	tggaacaaca	aacggaggag	gaggaggtgg	aggagaagga	ggtgggtgg	131700
gaacaaatgg	aatggaact	ggaacaacaa	acggaggagg	aggtggagga	gaaggaggtg	131760
gtggtggaac	aatggaat	ggttctggaa	caacaaacgg	aggaggaggt	ggaggagaag	131820
gaggtgggtg	tggaacaaat	ggaggaggaa	atggaatgg	aggaggaaat	ggaaatggaa	131880
atggaaatgg	aggggatata	gacacagacg	attttgagcc	tacgccagcc	cttctaaaaag	131940
aacgtctact	taattccata	tcttcaaaac	ctaaagaata	ctatgaagca	ttcgtatctg	132000
cagaggtaga	aactgcctta	caattatctc	gagatgattc	tacccaaact	ataataattg	132060
atgacgacca	attagaattg	gacgcctcag	acactctaca	aggaaaacca	agggattatt	132120
tgttcaagct	cgcaggtggt	tcacgcgcct	ttttagaagg	taccactatc	cgcaaagcgg	132180
aagaccgtgc	ccgtaatat	aatgaggaag	aaattgcaca	aacaatatta	agtcaactaa	132240
gagaaaaaca	catcaacgat	gaatacgaag	gaaaatatgc	cacacccgag	gaaagagctg	132300
atttttccaa	tagtcttaat	cttggtacta	aatacacaaa	ccatgaagtg	ggtcttttag	132360
taggagaaac	tattgaaaag	gctttccctc	atgagattga	gtttgagaga	tgcataattc	132420
tagtagagga	ttttaatagt	ggaactatta	cttcaaacac	tatgcagtag	aggtccaacg	132480
cttacaaaat	cagagtagta	gaaggatcaa	caacagatcc	aggggaagtt	gtcctgatg	132540
attgtttggt	ttttgccgta	gtagtaaata	aggaacaaca	ttctctagaa	atatctgcaa	132600
ctaacagatg	ccaagacata	tgctttgtta	ttattcctcg	tttatccgca	ataggaaaaa	132660
atgctaccat	ggtaataaag	aaaggtgatg	aaattaaaca	agaaacctat	ctgtttgtgg	132720
ccaataagaa	tgacactact	catttttcaa	tcacacaga	caaggatgaa	tctgttgtaa	132780
tgaattaaa	catgctcatt	ttctcagaaa	gaattctccc	cactttgagt	gaccctgcaa	132840
ccgttccaa	acctttgact	gacgccaacg	tactgtcagc	ctacggaaag	cgcctagggtg	132900
ttggtgcctt	tacagacaaa	aatttattgt	ccagccaata	aaaaaaatat	aaaaaaagat	132960
tgattgtttt	attttcaaaa	aaaataaaaac	accatgggag	ataagcaaaa	ggtggaacag	133020
cttttaagg	aattaaaggc	agaggcta	gatgattggc	tatcggtcaa	tggtgatcca	133080
attgttgaga	gattttgtcac	aactaaatct	gatgagacag	ccaagttgt	taaacaggcc	133140
gttgatgaaa	ataacgatga	atattagaa	gataaggttg	aagaaatgag	accggatata	133200
atcaatgaag	catccgaaac	atatgataaa	cttgctgtcg	atatgataag	agaggtagac	133260
actagtagtg	ttattgctcc	tgcaatagct	ggcacagtgg	caagaactat	caataattta	133320
agagataaaa	ggaaagaata	tgaaaagagg	ctatggacat	tagcctacaa	accatggaga	133380
agatagtac	aagcaattac	agtgatggaa	tttcgtttat	catataaaga	cctgactgtc	133440
cagccaatt	ccgatacata	tttaacattc	ccttttttaa	gaataaaaaa	gatcgcttat	133500
attaataacg	accgcgccag	ccccgtgaac	tgctccttgt	cagtggtctta	tcccaacaag	133560
tccgagtg	gaaatgataa	cggagttgga	cgtaaggttg	acatacacat	cagaagaaat	133620
gatttacag	aaaaggatct	ttacctgtct	gtttattgta	tgtagataac	tgatttttca	133680

ggatatgata	aagctgtgga	agttgatgcc	cataaatttc	atthttgaagc	aggggaacagg	133740
actatgtttt	tacccaaaac	ctctaatacta	tttaatcgat	cgcataatagt	aaactctaaa	133800
atatgacta	tagtggtccc	tcccgccttct	gccagttctg	cttctacaac	agaactggat	133860
aatgtctatt	acaggatcac	atgcacatgc	tcctagaggg	gggggggtggc	tgtgtccttg	133920
tataaaaaag	gaggttggca	cagactatca	aacattcagc	cgtaaaccat	ggagggaagaa	133980
agtcaacgag	tacagaggag	gatcggagtc	ctcccagaag	aggctgcctc	ccagatcctg	134040
aaagacacca	agttaagagt	atcatatctg	ggcgtgggac	attggggata	ctctgtttca	134100
gttataaaaa	gtgctctcca	gaaaggatgc	aggaggaacg	atgaggatat	aactgcgtgg	134160
tcaatcaggg	aggcttatct	ctattaccat	cttggtattga	attacattga	aaacgtaaaa	134220
cctgctgcaa	aatcattaaa	tacaaatatg	gtaaacagaa	ttaaaatcat	agctgtggag	134280
gataccagtc	ctagaagtat	ggtggcttct	aatgagtgtg	tgagaactct	agaaaaatat	134340
gaaaagggga	atthttaggca	acccagttac	ttgatggatg	ctgccatgag	gctagtccac	134400
gcctccagtt	ctagagtgtg	ctcccacatg	agagctttgt	gttgcaagga	agaggacagt	134460
gataaattag	gggggtattta	ttatgctaata	tttaatgaac	ttgaaactca	gtgtgttagt	134520
gcagtttaatt	tttctcccat	agaaaagaatt	aaacacgtct	tcaggggagat	tgaaagtgtat	134580
aaattggggg	agaaaagtgt	acagttatta	aatttaagaa	gtgtagcagc	ttaccatgtg	134640
ttaagataatt	atgggggataa	agtgaagat	acaaataaga	aacatagtgg	accattcaag	134700
cgaaaagagt	ttgagcagtt	ttgggggtta	tgcttttaaat	ttgttactca	gcacgtaaaa	134760
acagaccctg	aattacgttg	ttatttttaata	gagttgacat	atgccataaaa	ttggagaagg	134820
gattttttct	gtcctaaaagg	gttcttttaga	gaagaaagtc	tattccttac	atctattgtg	134880
gaattaattta	tagccatgtg	cataggtgac	cgtaaacagt	ttgccaaagt	tcagaaaagg	134940
gatttaaaac	gtttcaataa	aggagaagaa	gggagaaaag	aagaggctgc	tacatttgat	135000
tgatagagaag	gacatgttaa	acggatgcct	caaatgcctg	tatgggttct	ggacaaacat	135060
acgaacaaaa	acacacatgg	agtatctttt	gccctagaga	gtagtatggg	ttctggaggga	135120
gacaagcgct	gggtcccagg	ggtgtggctt	cattcttata	caaagatgcg	tctagattct	135180
ccccctcccc	cagaagtggg	ccaatttctg	gatacggctt	ttaatacatt	gaagcgagaa	135240
gctgctagtc	attgcgtaac	gaggaattatt	tgtactacta	caggatttat	aaaggcttct	135300
agttttactg	ctaataattaa	ttctgaaccg	atggaaatta	aggaagaaat	aaagaaaaga	135360
aaaatttgaga	ttaaggatga	taatactact	gctactgtta	ctgttagtgc	tactactagt	135420
agttctatta	cttctactcc	acctcctaca	aagaaacaaa	aaacaactcc	aagtgggaagc	135480
aataaagtag	actctataca	attgaataat	ctaccaactt	taaatatgga	ggatttagat	135540
agagtactag	aagtacacaa	ccaaaactct	aaaaagggtg	tagctgctac	agttttaatg	135600
aaggatggaa	ataaagttgt	gtttaaggag	atgagaaaaa	gctttgggtg	gggttctcat	135660
cagaattttg	ttcaagtact	aaaggatgaa	gatgtgtgta	aattggacta	tttgttgcca	135720
tgccctgata	gcggaccata	cagaggatta	tacagatgct	atttcaaaaat	agtaaaggat	135780
gaaatttcga	gtacagctgc	taggatagaa	aaagtgaat	ggggagaaaa	tgccatgtgt	135840
tatttcatct	ctgggtgtgt	gactagacaa	gaaggaattg	ggaaaattat	aacagatgtg	135900
cgcttttcac	atatgggacc	aaataaacaa	tatgtgtatg	ataactatag	gcaattaata	135960
cacattttta	tttttagatt	gttgactggg	gtgagtata	ccaacacatc	aaacattctc	136020
gtaggtgatg	gaggaaactt	gttctctgtg	gacgagaatt	atgtgggctg	aaaggatccg	136080
aggacagctc	tcgaaaatag	aaagataaag	gaactacaat	tgctactgaa	gacttctctt	136140
aaagtaaaaa	aggttaacaaa	ggaagatata	gtagctgtgc	ttccatcttg	gctatttgat	136200
acatcaaaaa	gtgataaaat	aatgaatggg	gtatgtaata	ttggtaaaaa	tatgggaatt	136260
ggtcccacta	ctttggatat	tgtaaaagaat	aattgtactt	gtatttttagg	ggctcgtgaat	136320
gacctgcttt	atgacaacaa	ataaaaggat	aatactttat	acaatttttc	atttccttcc	136380
tttttaaaagt	gtataaaaatt	tttctaagac	aaaagtgaat	ttgagggttg	gtcataatct	136440
agcacatttc	atcccttttt	ccccttcttc	ttcttcttct	tttcttctct	cttatttgaa	136500
agaccctggg	taggtccctg	ctaagggtag	gctgtaacct	accatataca	ctagagaatt	136560
ttactcttcc	tccatctcaa	aaacttttaa	aaaatttttc	tggtgctactc	gagtttagag	136620
ggtggaccgc	tggttcggcc	tgatgtcaag	ttcggagggt	ggaccgctgg	gtcgggcca	136680
tgtcagattg	caccagaaac	gtctgttgct	ccagaaacgt	acattaaatt	tctactgttt	136740
ctggaacag	cgtttctgga	cacaccagca	cccagaaagg	gtaggctgta	acctaccata	136800
tacactagag	aatttttact	cttcttccat	ctcaaaaact	tttaaaaaat	ttttctgggt	136860
cactcgagtt	tagagggtgg	accgctggct	aggcctgatg	tcaagttcga	aggggtggacc	136920
gctggctcgg	gccaatgtca	gattgcacca	gaaacgtctg	ttgctccaga	aacgtacatt	136980
aaatttctac	tgtttctgga	acaggcgttt	ctggacacac	cagcaccacag	aaagggtagg	137040
ctgtaacctta	ccatatcac	tagagaattt	ttactcttcc	tccatctcaa	aaacttttaa	137100
aaaatttttc	tgggtcactc	gagtttagag	ggtggaccgc	tggttaggcc	tgatgtcaag	137160
ttcggagggt	ggacctctgg	gtcgggccaa	gtgcagattg	caccagaaac	attgcgtact	137220
actcccatgt	ccagaaatat	ctagatcctt	tctggttctc	gcccaccacc	agatatctgg	137280
gaagccttag	gcgagtcagt	tttcttgcac	catgtatttc	taggaatttc	gctgacgtca	137340

```

ataaactggt ttgtcatcct gttttatcca gtttgctgac gtcaatagac catatttggg 137400
catctcgccc acacagaaac aggatatgat atcatagatt tctgggaaga ggggtgtttt 137460
tgtgccgatc tgggtgtata aaagagcgct gcggagagggc agaaacatca gacagacttg 137520
atctgtacag ctagcagcag cagtagcagc agccaagaga agatcggacg caaaccattc 137580
tcgcagccat ggcagcagca gcagtcacg gagagggag aatctctgca gatctactcc 137640
tgttggaaca actcaccgcc gacggagacg tgataagata cgactctgag cagtacacca 137700
aacctaggaa gatctttggt gacaaaagtg tgatagagac tattggacat tttctcatcc 137760
acaaccacaa ccaaggtgag agttacacaa tgcgcatcttc tgtgttgga aaattccccg 137820
ctctactcaa ttgcatatgg aatggagagt cgggaggaat ggctctatgg aaggcattgt 137880
acagggctaa aaagtataga cttctcaact cacttttggg ccacaagata aaaaattggc 137940
cttcggttgc cgtgatccct atctacgggt cagtgtgtga cagagaagaa aggccatca 138000
tcatgagtga gattattgac aaggaaactc ttcagaccat atgtaagagt gatatacgtt 138060
ctcttcttgg aatgatgaac gccaaagcat gcacattggg aggtaatttt ttacacttct 138120
atgcccgctc aactaaaccg tttgaaaatt tccaatatga agcaatggga gctaattgcag 138180
tgctaattggc agctgaagct atttatgatg gattcagaga ccatggctta aaccatcag 138240
aatatacttt tctgtgttg gaatctgctg atgtgtacgg aaacaatcca gtggaatttg 138300
caatatcagg agatgatgac aatatgttgt tgaacctcat ctgcaactat ggtgtatctt 138360
atgaaaaaac tcgcggtcgg gtaaatagat ccttgtttaga ttttttaaaa atgaacacag 138420
cttcaaagtg tctcagtggt ttaaagtttg ttgaaaaaca ctttaaaatt gagtcaaata 138480
caccaaaagg agaatttgaa gaaaaggcgg aaacatgtgt taattgtctt gatagaaata 138540
atgtgttgac taaaggtctt gaacaagaat cttacaagct ttcttgtgga cactttcttc 138600
atgtaaaatg tttgaggaat atttgtatag tatcacaaca cctaagatgt gaaaaatgtc 138660
taaaaagatt tgatgagagt attttgagaa agtgtacacc taacttgaat tgggtggtga 138720
ctatgccggc aggtgctgga aatgaagaag aaatatgttt catgagaaat aagaaactgg 138780
ttgatgattt cagaaaattg ttgtccccct tctcaattcc tcatttcttc aaaaatagta 138840
gacagcgaaa tcttgatatg ttgtgcccct acagtaccac cacgataata ccaataaag 138900
aagatccaaa gaaaaacgaa gatggaaaca gagtggggt caaccacaca gccatcagtg 138960
aaaagcagaa caaggaggaa gaagacgcga ggataaagcg tgtagccgtc aggacattta 139020
cagccatcag agaaaagcag aacaaggagg aagaagacgc gaggatcaag cgtgcagtcg 139080
acatggctgt cgcagccatc aacgaaaaga acaaggagga agaagacgcg aggatcaagc 139140
gtgcagtcga catggctgtc gcagccatca acgaaaaata caaggaggaa gaagacgcga 139200
ggatcaagcg tgcagtcgac atggcggttg cagccatcaa cgaaaagaac aaggaggaa 139260
aagacgcgag gatcaagcgt gcagtcgaca tggctgtcgc agccatcaac gaaaataaca 139320
aggaggaaga agacgcgagg atcaagcgtg cagtcgacat ggcggttgca gccatcaacg 139380
aaaataacaa ggaggaagaa gacgcgagga tcaagcgtgc agtcgacatg gcggttgca 139440
ccaccaacga aaagaacaag aaggaagaag acgcgaggat caagcgtata attgacttga 139500
ctgttgatat cgtatagtcg acatggcaat tgcagctgcc actaaaagg 139560
acaagaaaga agaagagaaa aggacaaaaa ggaacaaga gttaagggt gatctgagaa 139620
gggcaatgga tatggtgaac gaagtacaga agaaacttga agacatggaa ctgaaaagg 139680
ggtgtaataa ggatgaagcc aagaatacta gtaatgttgt tagtagcagt agtgtgttg 139740
cctattctaa agaaattgta ccttgtttag gaaataataa taatgctgtc attggtatga 139800
ctagcaccia ctattctgcc aacaatacta agaataatgt atttggttca cctcataaat 139860
tttcttcaa cgatgcattc agattctcca atattgtaga aactcccaa atgtctttca 139920
atttctcgtt caagacataa aataaaatgt gtaacatgca atactgttt ttgtaaaacc 139980
cttgtgtctg gataaaatc cacacatctg tgaacatttc ttgcacacgt aacactacga 140040
cttttcttca actcacacat ctattttttg cacatgcaag taaaacattt cttcaaacct 140100
ctacagagag atggtatata aaggccttgt atctccatca ttatcacta ttcgtacgat 140160
gacgtccaac agaccaacaa catccccctc ttcgttctct gaaggcttct cgctttcagg 140220
ggataagtat gatacgtatg aggatatctt acttgaacaa ttcaactgtt tcaagacatc 140280
ttctccttct tctgctcgta aaagtgaat agaggataaa actttaatct ttcaacttaa 140340
agaaggagaa aaattccatc ttgcaaaggg catagaagag ctccgtgaga ttctagacga 140400
taattctgca acaattgaac ctattatttc tccaacgaca ttcaatgaca gaaacgaatt 140460
actaaaccac gagggagata tatcctcaag tcccctatat actcagataa tgaagcatat 140520
ttcaccagag catgatattt atgaattgga ccttattgtt ggcactgatt tgcttttttg 140580
tctaggtgtg aatctacgca acgtttctaa actgatgaag aaaatatcgt atgggtactt 140640
aaatgtagtt gatgtgtgcc acagaaaatt tttcaacaat aggtattatag ttaatcccat 140700
ttctctgtca ttctcaaaaa atgtgtgtat tattcctcta ttttctgcag ctgaagaatt 140760
ttcgtctctg ggggaatgca gggatttatt caaccgtatt tgtgatgacg tagagagata 140820
tatcaactct ttttttttct accctgaaaa tactactact actactacta ctgctccttc 140880
gtcgcgccga atggaaattg cagatgagga agaacaatcc ccaaaaacta taaagagaaa 140940
tgacaacgca agtagaaact ggtctggtgt ctgtttgatt tttgaagtat ttaaaaacac 141000

```

gtactacatt	attaatagag	gagatagagg	aggttctttt	gaaaaggctg	tgaagagtgc	141060
aattttcttct	atcaaggaaa	agagatgtaa	aatcacagat	attaatggta	ataaacctcg	141120
attggttatg	gtgataactg	ggtgtttatc	agaattgtac	ttcaaagatg	cacttaaaca	141180
gattggagaa	aacaggcgca	aatttttgaa	aatgaatggg	aattactttt	ctctgattga	141240
tgaacaagca	gatctaactg	aattcgcgat	gagtgtttct	ggtgccgggg	agaggatttt	141300
tgtaaacggt	ttggggatgt	tccagaaccg	taaaatgata	cctgtaattg	atcctctcac	141360
atatgaaaat	gttgatgtg	gtgagcatga	tatacaaaaa	gaagatgcta	ttctttctgt	141420
aaggagagct	attgcagact	ataatgactt	tgtaagtaag	aacaagagag	ggaagaaacg	141480
cagcgcagaa	gaagaaaatg	aagatgaaga	tgcagacgct	agcagcagca	gcagcagcag	141540
tcctctcct	tcctctcctc	ctgcacataa	aaaatcacgt	cttcggatg	aaggcgaaaa	141600
atgtacactc	tgtaattttt	ttcaaaacaat	aaactaacca	ccttgatat	ttttgtttgt	141660
ttgcaacacc	ctttttatga	taccattaag	acacgatggc	ggtaaacttg	gataatgttc	141720
ttgtgaatat	caacaacaag	gatgaagatc	ttacaaaact	cgtatccgag	gcaataaagc	141780
ggcgagctaa	aactgtattt	gacactaaaa	atcaagcagg	gtttgacatg	agacgtcaag	141840
ttgaagctgc	attatatgaa	gcaatatcca	aaaagaaaga	aaaggccata	aaggcattcg	141900
atgagctcat	acaagaaaga	ggtgatgaaa	ttcacacttt	gactacaatg	cagtatgaag	141960
agtgggtaaa	ccgtacaata	actccctcat	tgacgactga	aaattttatta	ggtgatgttg	142020
agcacgccga	ttttttactg	gaccgaatga	caccgcgaag	cgaggaagat	attgaagggt	142080
tcgctgcttc	tacttttaag	gaggtatcag	attcaaaaaac	tgcaacagtc	atagtttaagg	142140
cgattgtga	aacgggggat	atcgatgaag	tgtataatct	tgaccatca	ttcggcgta	142200
ctcaagaaat	taaaatatat	aggtaaaaca	attcctcgga	attggataat	gtcgcagatt	142260
ctttccatat	ttataaaaatt	tctgcaacag	atagcgacag	tggcaatact	aaaaaattgt	142320
tgtatgggtt	aagggaataaa	aaagcagggt	atacgtgttt	gtgtagaatt	tttgcaaaaa	142380
ttgaatcaga	tgggattatg	gccaatacaa	atatcggtgt	cgctgaaaac	aacagagatg	142440
aaattgatga	aaacgaagaa	ggtaaatatg	gttttttaat	acccaaacaa	ccagctggtg	142500
caaaattgat	catctacttc	tttttaaat	gttgacata	aaaaatgata	ttttatacaa	142560
tgcaaccctt	ccttgggttt	ctggtctttt	ctgtgcttat	tgtaattgta	atgactgtat	142620
tagctgtata	caactgcacct	caaataaaga	aatcgaagaa	gagaaaaatt	gaagatgaaa	142680
atgaggaaga	accgcgtaag	actttggaag	attttgtaaa	gggtcggctc	cttaacgctg	142740
tcaaggaaaa	acctgcagag	tactttgagt	tgctaataatc	tgacagacact	gaagcagcat	142800
taaaaactgc	cgaagaaaca	gcccttcgag	attttgttat	tgagaacgac	tctgtcgaaa	142860
tagatgtgga	ggaagtactt	gaagagaaac	caagagaata	tgtcttcaaa	ttggcaggcg	142920
caacaagcga	aacgctaaca	aacacaatca	tcgcagaggt	acaaaaaaag	gcagcattaa	142980
taacagaaga	agatatcact	attaaaaatg	taaaacaatt	cagggctgcg	aacaaagata	143040
ataaagacgg	ggaagcaact	cctgaagaaa	aggaagattt	taccaataat	tcagatcttg	143100
tggggttgta	cttgaacgaa	gtagttagaa	aaacaacaaa	tattgtcatt	aacaaaatat	143160
tcctctcatga	gatggttttt	gaaagatgtg	ctatttttaat	tgaagatttt	gatactggtg	143220
ttgtgactga	tcaagccatt	cagataccct	ccaacaaata	caaaatcaga	ttagtccaag	143280
gggatgaacc	tgaagtattc	cctggtgact	gcttggtact	tgacgtttca	gttgataaaa	143340
taaaccacgt	cttgaaaatt	tctgcaaaga	acggatgtga	aaacaactgc	ttcgttatta	143400
ttccacgggt	ctctcctgta	ggaagtgttt	cttccatgat	attgggcagc	actgaccaag	143460
tcaaacctaa	aacattctta	tttttagcca	acaaaaatga	cagtaacacat	ttttacattca	143520
caatggataa	gcaacattct	gtagggtgtg	agttggacat	gttaattttt	tcagaaaagga	143580
acttgaggaa	tttaccogat	tcaaaacctt	gacctctaag	tgatgcagac	atattggcct	143640
catatgggaa	gcgtctagga	actggtgttt	tcacaacaga	aaatttggtt	gacgattaaa	143700
taaaatttat	aaccacttta	tctgtagact	tttatgtcct	ttgtgcaaca	aaaataaaca	143760
tgtcttcttc	gtcttctgaa	actcctaaga	cttccaccga	tactggggaa	gaaaggatta	143820
aagacattgt	aaatgctcta	gataataatg	gcgagtgtgt	gtcttcctat	attgatccga	143880
ttatcaataa	tcacatctca	cgaaaaacgg	cagaaactgt	ccaaaaaatc	aaccaagaag	143940
ttgatgaacg	gtacgataga	aaaatagccg	acaaaatcaa	cgaaataaaa	tcattccatct	144000
ttacaagtgc	tcagactatg	tatgaccaat	atgcaataga	cacatttcaa	gaaggaaaaag	144060
gagccaacgg	gtcgtggacc	gtcatggggc	cagtgaacac	ggttatcgat	acaactttaa	144120
ataaaaatgag	gggaaatatg	ctcgaatatg	ctgaagatat	gtgggacgga	gatgactgga	144180
aacgattttc	cagttctatg	acaacgcttg	aatttgatct	aagttactct	gatttaacta	144240
tgatgcgtgg	ttctgacggg	tattttgcat	tcctttccg	tggaacaaaa	aagataaaga	144300
tggacggttc	aagaaagaaa	gaagacccaa	ttaattgtat	catttcagta	acatatccaa	144360
acaaagtagg	ggatgagtgg	gaagagggtg	aagaacgtga	agtgaatttt	aacctagaaa	144420
gagtagacga	ctatgaaaga	gatattccatg	tttcaatttt	gtgcatgtta	catgcacaac	144480
ttgataattt	cgaacaagca	ttaggagaaa	atgcaaaactc	tttttatttt	aaaaagggggc	144540
aaagagtcac	gttcttacc	aagaaatcta	aactgttcaa	tagacctact	gtagaagatt	144600
ctgatatgtt	ttctataata	ttccacctg	catctgacca	agattttgca	gatgatattt	144660

attatcgaat aattgtaaca tgttcataat aaaatttgtt gaaataaaat acagtaatac 144720  
attcaaaatc ttttattatt tttcccttct aatccaacat atataatggt gatacattta 144780  
caggtgatac ctgaacatat ttataaaaaat caggaatata cataaacatg ctattgtaag 144840  
ttgtattcat atccctgggtg atcatctcgt gccagatgcc cctcagcgag atggagcttg 144900  
ctgtaatctt tgccctgtgg agtgcttgtt tgtatgcgtc ttgagccctc ttcatctcat 144960  
gtgctgtagt ttcaccccc tcacaaggct ccatgataat attctttctg gccgaattat 145020  
acttgtaatc aatgaggtat tgcttgaaag gccagtgcaa gatgtcgttg gttgtagtct 145080  
ggatgaggtc acatttcttg gcaacattct taaaggcggc agcagcagtc ttggcagaca 145140  
ttgaattcgc ttcaatgctt ccaaaaaat cctcattatt ctcgtgcacc ttcaagtagt 145200  
tgccactatt ttctccacg cctccgtcaa caataagggt gcgaacgtat ctgcggctgt 145260  
gtgtcttgtt ctcaacaact tccctcttta gaagatttcc aaccatctta gtgtacatgt 145320  
cagatctcca attgttgttg atgtcgttat gttgccaga aaagttccct ggcaactaa 145380  
ttctgaagggt cttgtagctt ccataaggag tagtgttatt aatgtggtta atgggcatag 145440  
attcacagaa tggaatagcc agtgatgga cagtagcgta taccttgga gtcacattat 145500  
cttcacatag agcattaata agagcactct tcttgctctc acaataaggc ttcacagact 145560  
tgtcgaaga ggaccagaaa ctcttatcat aatctacctc caccattgaa gattcgtatg 145620  
tagtggaacc catagaaaag ataataagat tgtccacatc gacaagtact gagtagtact 145680  
ggctgttttg cagatcaagt cctgatcct cctcatcttc atcctcatcc ttcttcttg 145740  
tagactcgat gtgcacacaa tcctcagcag cctcagttcc ggcagggaac aacttaagag 145800  
acttgaggat agtgagcatg ttgaatgagt tggtagtga tccagaggag tgtctgcata 145860  
gagatagaca gacaggagt gatctgttct ttgataggag agaaacatcc atctcttctc 145920  
cagattgagg atatttgggt actgaagcag atggtctgaa gctgttgtct gaagccaggt 145980  
agagaggggt cccatcgctc tcgtcgaatt cgcactctgt tacatcccca aacacaactt 146040  
ctgaacactg ggaagtgtca ccatcatcat cctcatcttc ttctgaatct tcaactgtac 146100  
tgctgataac actttccttg tcattctcat cctctactac tgttgctgta tccttatcct 146160  
tatcctcctc atcctcttcc tcttcttctt cagcagcagc tacaacagt acagatgaag 146220  
acgaagagga agaagaacac actgatgagg cgtcgtcatc atcatcatcg gactcttctt 146280  
cttcttctc tgactgctca cctcttctc tttctcctc tacatcactg agaccgatag 146340  
tttttcta atgattcttg atgacatttc caactgaact ttcaatagtt tcattgttgt 146400  
aatcgtaatc aacttctcc tcgttttcat cttcgtcttc atcctcactg tcgctgtatg 146460  
catatgctga cttttcgta tcttcttctc caaatacagt attgttcagt gacttgagg 146520  
tactctcct tttggccaac cacttctgt ctttgaagca acaacatcca ttctctatag 146580  
attcagtgac actttcttcc acaatttggg tagtggtctt gaaatgtgtt cctgccttct 146640  
tggtagacga agaactggga gttgacatga ccaatccagt tgacttgatt gcactatcaa 146700  
tatagtttcc tccaagtctc ttaaaccagt acttatctcc attacttcca aatatctgcc 146760  
tcttcttctc caaatccatc tcgtccacat atacagcatc agagagcttc aatgtgaaga 146820  
gagggcctcc cttatcaagg gccttcttgt tgagtgttct aatagtacag tgctcatgga 146880  
actggagatt catacaacca aggaatagat tggccagttt tgtagtatca gtggcgattg 146940  
cagttttctt cacagcttct cctcaacct tgacaatgtt gctaaccaga tcctttcctt 147000  
caaatgccga cttgacctt gagacacacc catcacaata ggtggaagg gacagttctg 147060  
tgcgctcaga gttcatcaca ctattgaaca ctaccgattt ctgcgccttc ttatttctt 147120  
ccaattcttc gtcctcatct tcgtcattct ccagaaattc agcaacagat ttggccagaa 147180  
cactcacgtc aagatgagct ggcaccttg ccataagagc gtccagttct tgaagagatt 147240  
gctgcttctt ctgcgggggt gtgatttctg gaggagtctg ctgctgctgt tcttcttct 147300  
ctcctacaac ttcttccaca gtttccgtgt gtggactcct gatagtgttc ctctctgct 147360  
tagagaccat tcttcttcca gcagcaaat tggccttgga aggagtgggt ttcttggcag 147420  
aagtagtagt agtatcctca aattcataaa cttgatctgg ttctcagag aaagtgcag 147480  
ttggctgctg ctgctgctgt tccttgattg actccatttt tagaatattg ggttattagg 147540  
attcagtagt tctgttatcc cttagacaaca cgtcttgtct gtacgattag tcgagatgat 147600  
gtgatgtaga gcccttgccc gtgcctcaat atataccct cagtagacct tgataatcac 147660  
ggttgccaaa caacaacggg tggaatctt agtagaaaga accaatggaa caagtcagga 147720  
taataatata cagtaataag ttttgggga caacattcat ttattgtcat agcgtaaca 147780  
ttagaagtac aagaacagg gattctgcat ccagctcatt ttgttcaacc cacttctatt 147840  
gcaggggaca acactcttct ttcttgaaaa atatggatga agacacttg aattcttcaa 147900  
taggacctgg ggcgttgctg gaggaggacg cagaagaaga agaagaacag ctgccgttg 147960  
acgagtttat agtttccaga agacaacata atttagctgt aattggacga ggagacttct 148020  
tgtaatgaga gaaccacaaa attcttcttc ccgtcttttc tggataaaaa gtagtcccat 148080  
aatatgttcc tttctggact accacattat cttcgtttaa ggaacatat ctgaccgact 148140  
ttctcatcaa gggctgatct ccatttcttg ttgtgcttag agccgatttg tattccccgt 148200  
cgttcacgtt atagttgcat atagaacagt gggatgaaag gaactgtctg gttacgtgtt 148260  
cagtttgtac aagataactg gacgcatttg ctagttctgg agggacgtcc tggtttttgt 148320

ttattgttga	agtaggtttt	atcttttttc	tatttaagca	cacacccttt	gaattaatgt	148380
gtatgcgatt	cgcattgatg	tatatgcaga	ttgtgagtgc	tggagaaaat	ggtatctcaa	148440
aatcggtagc	tggttcaata	atctctggtt	caacttcctg	ttgttcttct	tcctcatctt	148500
tctcttcctt	atttttgaca	cctagaactc	taaggttggc	agtactaatt	attctggtctg	148560
cagccgtgct	ggtagcagaa	ggggaagatt	gttgggagat	aacaggagac	atgttgaccg	148620
tgccgtgtcg	tgaggataaa	tggctaaagg	tatatacggc	atgccctttt	ttccgttact	148680
gtgtgcgagg	tcgttttcta	ggaccgtaga	tgaccatctg	ccttcctctc	tttttctgga	148740
caacctttta	ttgccatggt	gttatgagag	gtaaactcgc	ctcatagaat	ccgcaaactt	148800
cctctaaagg	gaacaagtgt	ttccccatag	gggacaattc	aaaaatataa	ttccaacata	148860
ttaaggggtg	tgtttttatt	tccttttcat	cgtactcctg	tattgtttta	ctaaaatttg	148920
gaaggtgatt	ttttgataga	ctatcggaca	ttatatccgc	tgcaagttga	cgtttatgtt	148980
caatgacata	atcattcgaa	tcaaagttgt	cgtgtgttga	caccatacca	ttccatttcc	149040
atgttgctag	tttgagaggg	gattttattgt	tgcgcaattt	tagacgtgca	gaaacgtttt	149100
gccatattat	tttagaaagt	acaggcaaaa	caacatccct	cacgtcaata	ctgtccatta	149160
atatatcagt	agaatcattg	aacctttata	catctctctg	tatatcttcg	tttgatgggg	149220
ccaatcgaac	tagaggaact	aaatccccag	ttatgagatc	ggtgaaaata	aagctctcgt	149280
catcaacaat	ttttttggag	cctatgtaca	tctcacactt	actataaaaag	tcatttttat	149340
accctggttt	cttttcccac	ctagaacata	agagtattaa	aactctcgac	aaattataca	149400
tggcacctac	tcttccttcc	atattactga	tagctctata	catattaacc	acgtattgta	149460
taggtgcttc	tacgtacaaa	tctgaaaatc	cactagcggc	agttccttca	ctaaaggcta	149520
gaacttgctt	catagtcttc	tcttctgaaa	acatttgctg	aattatttta	gtggatagtt	149580
tgtgcccgtt	attgaatttt	gtgccagggt	actgggagcc	gaattcgttg	aaatgtgtgt	149640
gttgtttttt	ccctacttct	ttagtttcat	gccaggcggg	gtgttcttga	agagcagcca	149700
cctccaggaa	caaatcacag	atggctcgtt	ctccatcaca	tgtgcatttt	acagtttcag	149760
ggaaatttaa	tcctgttatt	ttataaaaaca	actctctgtc	tattattttt	cccattcgtt	149820
ggcgatttatt	gacattaaac	ccacaacact	cgctctgat	tccattata	aaggctaaag	149880
aatatactat	aatatttccc	gcacttgaac	ccagcagctt	aagagttgta	tattctccat	149940
cgcagttata	gagacatctt	gccaattcgt	ccattgattt	cctgcccgtg	aagcgcacat	150000
ccatgagtct	tgctcctcca	tatccaattt	cagcgtttat	tgaagtcctc	caacctttat	150060
ccttgtttag	gcgctgtatc	tttttcgcct	tctctgccaa	tacattcata	aaagtgtaaa	150120
atttaacgtc	tcttacagat	tccataattt	tttctcaag	tgtagcggc	ggcgggtggc	150180
gcgatccttc	ttttagcctt	tcaaagataa	aagagtgtta	ttcaacttac	agtgtttatt	150240
ggcccagcaa	ttcttttctt	tcatcaactg	catagtagaa	ggtcgatcgg	aaagaatgac	150300
caattttgaa	tccaaatccc	taggtttatc	tagtccagct	tctatgtcca	tactctttcc	150360
tcctattatc	atatagtctt	caaacagtgt	tttattagca	gacatgttaa	tcatttcaat	150420
tacagtctgg	aggcatgttt	catgtctatt	gaccagatct	tcatcatctt	catcccatgt	150480
cacacccacc	agaattattt	ctcctttctt	gagtttggat	acattttctg	taactctttt	150540
cattatccgt	tgtttcttgg	acgccacact	ttcgtagcat	agcgtcggta	ggggagagtc	150600
gcaatccaca	accaagtatt	ttttcatgat	taatttgttg	aggatttgtt	tctttgaatt	150660
gaagaaaaaa	atcattcgtt	ctgtaccatg	gagacaacta	tggataatgt	cgttcagaat	150720
aacgacgtaa	caaaaccaac	accagatgtt	gctactgtta	caacagcaac	tgagaaacgt	150780
cagtcatgca	aagagaaaaa	ggatcaactt	aaggccgaat	gtcctcaagt	actgagagca	150840
ctaaaattgt	ccaatacttt	aaaggcaaat	tttggaaaaa	ccatgtcggc	tatttttgct	150900
caacatttag	tggacatgac	aaacgctaaa	cactttaagg	acccaaagac	aaagaagatt	150960
ttagaactgg	atggaagtag	tagcagtgac	agtgaagaag	aggaagaaac	tagttcttca	151020
tccaaacgga	aaagaggtag	tgggtgctaga	agtgtctctt	caaagaaaga	aaaatgcccc	151080
aatactatca	aaaattggct	caatgatgct	caaggtgtat	tccgccagtt	tgcatatc	151140
atcattaatc	ttcctctctt	tgatgatctt	agagacgaag	taaaggatga	acaaactgag	151200
ctaaagacca	tatatgactt	gtatagacag	gacatggaaa	aggtggtgga	agaagtttta	151260
gggcgccaag	acctgtttga	tcacaagtca	gaaaatagcca	aaggttttggc	ccgtttcgat	151320
accacgtctt	cgttgcctcc	ttcggaatag	tcggctgttc	tagactcgtc	catatccaaa	151380
aggttggaaa	aaaatagcaa	gggcccgaac	agtaaatatt	ttgacacact	aaacacactc	151440
aaggaagaaa	tcaaagaact	tttatgtcat	catgtcaaat	atttattgca	aaatcttaca	151500
ccggaggatg	caaattttgt	gttcaatagt	tctgtaaagt	atgtaaagaa	atcatatcaa	151560
tactacatac	aaacatcaga	gatggaaagt	gatgaattta	agtcccttct	cactggagtc	151620
aacattaaaa	tattggagaa	gataatctcg	tcgataata	atgttgctac	tccttacaaa	151680
cacatcacta	atcccaggaa	cattatttctg	tctttacaaa	aagtacgtga	aactaaacct	151740
gtttcaaaag	attatccgtt	cagagtggat	acggccagag	atattgtact	acttccagag	151800
acgggtggca	ttttctgatct	ccctatcaag	ccgtttacat	tattgcagtt	ggtgtcctac	151860
atcaacgctc	tcttttccct	cgagcgccga	aatgttttca	ccgacggctt	ttttaatgca	151920
gcgtgcgtcc	taatttccca	gtgcctaacg	aatgcgaacc	tattatctaa	cgactttcct	151980

aaacccattg	aactggcagc	taatgttact	cgccataatc	ttctgagtat	gaaaatgctt	152040
caagaaggtt	catccagtga	aaagaagagt	aaaaagaagg	agaaaaagaa	ggataaaaaag	152100
aagggcggtg	gtgggtggtg	cgatttctgat	tcagaaacag	attcttccctc	atcatcatca	152160
tcttcttctt	cttcatcctc	ctcctcctca	tccgaagacg	aagaagaaga	aaaaggagaa	152220
gcagtagaaa	aaaggcaagaa	aactaaacgc	aaaacaaaaa	agaaaccatc	aaaggacgac	152280
gatttagata	caattagtaa	actgattcta	aaaacaggag	gttacttcca	cgacacgagt	152340
gaactcggca	ataaaaattag	aaatttaata	gacaaggatg	atthttgcggg	cgtagcccaa	152400
tatgcagtaa	caatcactga	gatgcaatct	acgccaatga	atcaaagatt	agtatctagt	152460
cttttagatt	tgataatgag	actaaaagaa	caagtaaaat	atagtgttga	tactgaaagt	152520
acttctcca	ccgccaaatc	taataatgct	ttagatagt	ctaaattgac	atctcaacaa	152580
gtggtttacaa	tgatgggtgga	ttctggagct	gaattggcaa	ggctagctgc	cttctttttc	152640
gtggtggtg	ataatactgt	agttaatcgc	catgaagcat	tcatttctaac	atcaaaactt	152700
ttaccctcga	atgaaaatag	aggacttaaa	actgttgtag	agtcattctt	taaaaattta	152760
acaattagta	acaaagtctc	tacctcgaat	gaagaaatga	tgtcgggtgat	gccgttcgaa	152820
gacgaacaac	aacaacaaca	atgcctcaa	catgaacaac	aaccggattt	gaaaagagt	152880
gtgggagaag	tatttctaga	aatgggaaaa	tcaatagtga	actcattccc	ttccaataag	152940
agtgtacaat	taacggctga	cgcggttcaag	caaaactact	cgcttatggg	aagacgcata	153000
aattttggcg	ccaagataaa	aacggctata	tctatcggat	ctaatatctc	gcccaatata	153060
ctattttcta	acctcccaga	atctgtaggg	aataatactg	taactggtct	aaggctaacc	153120
aacctattga	aaaacatata	ccaaagtgcc	caagctaata	atataatcaa	aaatgccaac	153180
acgctcgtaa	acaatacgat	ggaccaacaa	aactcagcag	ccatgtccat	actactcttt	153240
cctccaacat	caaaagaaac	ttctatcttt	cccggaaacg	atccttcac	tataaaatta	153300
caggatatga	ctacaatgtc	aaatctggca	cgaggatttt	attccatcgc	agaaggatgt	153360
atcggggctg	tacgctcgag	ggaatttgat	gaaggaggag	ttaaagcgta	cactttacta	153420
gtggactcaa	atactatgga	catggcagtt	aattttgccc	ctcagtcctc	ggaaaaatca	153480
atgtctgaag	ccttaacaaa	taatgcgaat	atgaaccctt	ctaattgtatt	agaaggagga	153540
tctttttaga	acgggtgctc	ttcttactatg	tttgaaaaaa	atggcagtga	ctgtgaaccc	153600
acccctctag	caaaatatac	aatgaaagat	gtatcaaata	gatacttgaa	aaaattcaac	153660
aatgataaaa	atacacaaga	tttatataaa	aatagagcag	aaagagcact	cgtcgaacag	153720
gttacaaaata	agcctacaag	tgttggtgcat	tcccagctag	caaacgccat	gggggtagct	153780
gttataggcg	cagcatcaat	caagctcatg	gaagcagaag	cagcggaaac	tgaatgaga	153840
gcagcaaat	atcaagcaac	atcaaaatct	acaaatgcta	ttaataataac	aaatacaatt	153900
ggaatgatac	gtaatactac	tcacctgtgt	acgaccatcg	ccgtgagtgc	agccgccgac	153960
atgtcaaagc	tcgccaacaa	ccattttatg	agtgtattaa	acactgcaaa	taacagtcac	154020
agcagcagac	gaggagacag	atcgctccctg	atggtgcagc	aacaacaacc	aacacattca	154080
gcattttctg	aacagacaag	aggaagagga	ggaggagtat	taggatcagg	aactgaacag	154140
acaaaggatc	atgtggaacg	tatgaagaga	gattggatat	taaacaatgat	atctccagaa	154200
gacaaaaata	ctactactac	tacacccagt	aatgccggcc	ggacattagg	atatggatct	154260
aatataactg	gcataaacac	gataaaacaa	gacgacaaga	gtatgatgga	taaactttct	154320
gagatgtcta	gcttcagaac	ttaaagtgcc	aattttcata	taaaaatgta	gtattatgtg	154380
gtcgttttta	aacttttaagc	gtttgtggtt	ttgggtcggt	cacgggtccaa	gatgtaagag	154440
tccccttgta	cccaaatatt	aacaaagtgt	cagtatggtc	aatggaaaagt	atctaaaaat	154500
ccaccctgtc	atcattcaat	atcccaactg	aaccaaccaa	ccccagaacc	aacaacatgt	154560
ctcttggtga	gaataatacc	caagaagaaa	tgattctgga	aactactggt	gaagggtgtg	154620
tggaggggagc	agaagttgcc	cctagagggtg	tcaagagacc	ccttccctct	tcctcatctt	154680
cttcttcagc	ttcggatagt	gaagatgatg	aaggaggaga	gcaaccacaa	acgaaacccc	154740
caaagaagaa	acgtaatatc	aatagcggaa	agtactggaa	aatcgaaaca	atcgaaaccag	154800
catctccaga	aatgctcagt	gctgtaaatg	atattgataa	cgtgtccaaa	actattcctc	154860
taatcgataa	tagttttggc	gttcagttta	aaaagagtgt	atctgaagaa	cagatcaaaa	154920
cgctactcac	tgaactattt	gccgtggaat	atggaacaat	cacaaatggt	aaatattcaa	154980
cctttaatca	actcgaaagg	accggagagc	ctcttaagaa	gaagcgctca	aacaatggaa	155040
ataacaacta	cgataactgg	caaattcgaa	ttgaagccgc	tgacgcagag	aatgtttacac	155100
aagctgtttt	ggacgcaatc	gttgaaggaa	atgacaccgt	gatcaaggca	atctcctctc	155160
ctgaagggga	gggaatcggg	cttcaattta	acaagagtgt	tagctctcaa	caagctaaaa	155220
atattgtcca	ggctgcgat	attgaatttg	gacaagttgc	gcacatgaag	tgtaatttgt	155280
tccacaagat	ggagaaggcc	gatgaatctt	ctaattcttc	tggtgaatca	ccaaaggtta	155340
agaaagtaag	gaggaacaag	tcccagccta	caaattctta	ctacactttc	actatgatcg	155400
gggattctct	tcaagagcgt	attgataacg	caatcaaaat	tatcgaaatg	tctcagctta	155460
agaggccatt	ctccaattct	gctgctgctg	ctgaagaaga	caccactact	acgaccacta	155520
cttccactgg	tgttgttaat	ccacgaggga	ttgaaggatat	tcactttttc	gattcatcca	155580
tttctaaggg	atgttttact	gtgaggaata	ttgttgtagc	aaatgggtgaa	gttccacaag	155640

aagaattcgt	ttctgagttg	tacactaacc	ttctaaaggt	tgaagagaag	gtggatcate	155700
ctactttcaa	gaagcttatt	catgaccgta	caatgaatcg	acacatttaag	gcatgggtact	155760
gcatctgccc	ctactacacc	acaggaggcg	tccccctgc	agctgataag	gtttcccgcta	155820
aggggaattgc	tacctacaga	atatacgagg	acaggactgg	agtgttccaa	ttcgatgggg	155880
cacatacatc	tactaccca	gcacaggcag	ccgaggcaac	tggtgctatt	cacaagtcta	155940
tgctcttcca	gagtcctgga	actgatattc	aaaagttcct	cgatgctaag	aaggcagaag	156000
gtttggagcc	tatttcatcg	ggcgaaattg	tgtaccggtc	aaagtggagc	cccaatgata	156060
gcagggcaac	acgctgcttc	aaatttttatt	cctcttcaga	cgagaagatg	aacattgctg	156120
atgttttgtc	tattgttcat	actgatggac	ttttcagtag	tgtacacttt	agaaaggata	156180
ctatggaata	tggtgttgca	aagagcaaga	gtaaaattat	ccccaaaact	atcaagatta	156240
agaagggagg	agatactttc	catagcgagg	aggatattga	agtgcctgta	aaattcactg	156300
caatcacttc	ggagggaactt	aataggggaat	gtaacaccaa	gggaatgaac	agcctccgag	156360
cacacaagaa	acgaaagagc	aactcttcta	ctactactac	ttccaccact	tccacttcaa	156420
caacagcaaa	tactccaaag	aagacgaaga	agagcgcttc	tgctgccagt	gacccatttg	156480
caaagcttac	tttagattat	gttgatagta	catcatttgt	attttacaat	atcagtaagg	156540
aaatggtgca	gagaattttg	gctcaggaga	gagtaagac	gctaaaggca	gtcaagaacg	156600
aagagaagat	ggaaattgta	gagggagaag	aagcacaaga	aacttataga	ggaatcgtaa	156660
agatcaagac	gaatgcaaag	gcatacaatc	ttgccaacaa	aacatctgga	gtacttttcc	156720
ctgctgataa	ggtgtgtcta	aagcacacat	tggaagattt	gggggatgta	cttgattttg	156780
atgttgtaag	ggaggataat	gttaataaga	cggtcgcttc	tactactact	acatcatctg	156840
aaaataaggc	aagcggagga	gatgatgaag	aaactccaat	ggaatttgaa	actgacggag	156900
agaagctggt	gcacgaattg	ttgaatgaat	aatgttggtg	tggtgtattt	tgtatataga	156960
gaaaagaata	aaataaaaag	ttaatatata	gttgcttttt	acatttttct	aaacttgata	157020
taataaaacta	tatagcaaat	atlaatgcta	ttgacacatg	ttactagtcc	ctagtcaacg	157080
tgtaattgta	catagtaaca	tgccttttct	taggcatttg	aggcaaatag	gaaaagagta	157140
tttacgatgg	gtgaccgtcg	tgataacgcc	taccacaggt	gacagtatag	aggggtttcag	157200
cagactgcaa	atactcacct	aacaggcctt	ggtatagagg	aatcaataaa	aaatcaacaa	157260
aaaccttaaa	agtttggtgt	ttcatttttt	tacatcgggg	acagaaatgg	ggttaaaaac	157320
aaatatccga	gctgcaatct	gacattggct	ccctaccag	aggtccaccc	tccgaacttg	157380
acattcagtc	gacccagcgg	tccaccctct	aaactcgatc	gagctgaaaa	attttttgaa	157440
aagtttttga	gatggagaat	gagtaaaatt	ctctagcggg	aacagtaggt	tacacactac	157500
cctttctggg	tacgttcata	tccagaaaacg	cctgttccag	aaacagttaa	gagttaatgt	157560
atgtttctgg	agcaacagac	gtttctggtg	taatctgaca	tcgagtcgac	ccagaggtcc	157620
accctccgaa	cttgacattc	ggtcgaccca	gcggtccacc	ctctaaactc	gatcgagctg	157680
aaaaattttt	tgaaaagttt	ttgaggagga	gagaggaggt	aaattgatac	agatagaggt	157740
atcttttagta	gaggtggtgg	tccttctagg	agctggtgtc	tccagaaacg	cctgttccag	157800
aaacagtcac	aagttaatgc	acatttctgg	agctatccca	tttctggtgc	aatctgacat	157860
tggtccctacc	cagcgggtcca	ccctccgaac	ttgacattcg	gtcgacccag	cgggtccaccc	157920
tctaaactcg	agtgacccag	aaaaaatttt	ataaaatttt	gagtcggaga	tagaagggaa	157980
attctctagt	gaatacacga	ggttacacac	taccctttcg	gggcgcagca	agtctccaga	158040
aacgcctggt	ccagaaacag	tcaaaaagta	atgtgcggtt	ctggagcaac	aaccatttct	158100
ggtgtaatct	gacatcgagt	cgacccagag	gtccaccatc	cgaacttgac	attcggtcga	158160
cccagcggtc	caccctctaa	actcgagtga	cccagaaaaa	tttttataaa	attttctgta	158220
agggaaaaaca	acactaagtt	gcttgctaat	tcatacacta	gtgcaggttg	gttgagatag	158280
agagtgaagg	aggctatcca	gatattctgga	tcaaccagaa	tccagaaaacg	tttctatcca	158340
tttctggaac	agccatttct	ggaaaagggt	acatttttatt	atatttggtta	tattatttct	158400
ggtaccactt	ctggcacatt	cgtgcaatct	gacattggcc	ctaccagcgg	gtccaccctt	158460
cgaacttgac	atcagggcca	cccagcggtc	caccctctaa	actcgagtga	gctgaaaaaa	158520
tttttgaaaa	gtttttgaga	tggagaacga	gtaaaattct	ctagcgaaaa	cagaagggtta	158580
cacataccct	ttctggttac	ggtcatgtcc	agaaacgtct	gttccagaaa	cagtcaaaag	158640
ttaatgtacg	tttctggagc	aacagacggt	tctggtgcaa	tctgacattg	gcccagacca	158700
gcggtccacc	cttcgaactt	gacatcaggc	cgacccagcg	gtccaccctc	taaacctcga	158760
tgagctgaaa	aaattttttg	aaagtttttg	agatggagga	cgagtataat	tctctagcga	158820
aaacagaagg	ttacacctac	cctttctggg	cgcagctata	tccagaaaacg	tctgttccag	158880
aaacacacaa	aggttgatat	acgtttctgg	agcaacagcg	tgacagtagt	ttctactaag	158940
ccttcacttg	tacacatttc	acttagtata	atttgtcatc	acatggctga	caaaaaattt	159000
ataaaaaattt	ttctgggtcg	ctcgagtta	gaggggtggac	cgctcgctcg	gtctgatgtc	159060
aagtctggag	ggtggaccgc	tggtgtcgact	cgatggcaga	ttgcaccaga	agttgtgtgt	159120
gctccagaaa	tggtggggga	gtattcagag	tgtgtttctg	gaacaggtac	acatacccaa	159180
ctgatcagat	aaaaattggg	tgtatacatt	ttctgatccg	agaaagtaaa	ccaactgaac	159240
ggtgtaaaac	gtactgaaac	attacatcac	ccacagtcgt	cctcctccct	tctaagagag	159300



aacatatccc	gtacccaact	acctcaaaca	acaacaacaa	caacaacaac	catgacatcc	159360
attatggacg	atctctacat	gagctgttct	tcttcttcat	cgtgctcaag	ttcatctgaa	159420
gaggagaatg	aggtaggagt	agaaggagga	ggaggcagaa	tcggaccaac	agaggccaag	159480
aaaaagatcc	tccgtaaacg	aaagagatct	tctgttaaaa	gtacatcatc	ttcttcttcc	159540
tcatcatcat	catcagacga	ttctgactct	gatagagaag	aaaaagaagg	aagaaaacta	159600
tacgtggata	ttgcagatac	aaggaaaccg	cccaaagtaa	gaaaactgga	tactccttca	159660
caaacttttag	agaacgacct	ctacatgtct	agttcatctt	cttcttcttc	ctcatcttcc	159720
gactcatcat	cgtcttcttg	tgaagaagaa	agtgacgatg	atgatgatga	cgattatgac	159780
ccagataatg	tccatgttct	tggttgtaag	aaggaaaaat	ctccccaaga	catagaagct	159840
gaaaaggaaa	aggaggaaga	gtatgaagaa	gaattcaaga	gaatggcatt	accttcacgg	159900
ataaacacat	ctgtagatga	ttgtgttata	cctgatcgga	ttttaaccct	cttttctaca	159960
cttttgaaga	aaaatagttt	ccagtctctt	cagccagttt	ctttcctccg	gttgggtgatg	160020
aagcaagtaa	acgaggccat	gaattcagca	ttctcttcca	tgttatccag	ttctgggtatg	160080
aggttgggtg	aggactcttt	gggtgacaca	tctaaaattt	cctcctttat	aacacctcaa	160140
acggatacta	gtaattcatc	ttcatcatcg	acgttctgta	ataattgtac	agatgaagat	160200
atcaagaaagc	gtaacattgc	aatgggaaga	gtagcagaat	tgctgtcaaa	tattgcagcc	160260
agttctaata	aggagaataa	tttccgccct	gtagtttctt	taatgcgagg	accaacttgt	160320
ggaggttcta	atgcatccaa	caagaaactc	aatagtaaca	ggcaaactat	tccccaagta	160380
ttaaacaagg	ttatatTTTT	tagagaaata	cacagtgtta	tcgcgctata	tttatcgctc	160440
gtctgtgtgc	agcgagcaat	gaataatgac	aacacaaatt	ctagcggata	cgctgaaggg	160500
atggttacta	aaatcttgaa	tattattggt	aaaattcctt	ataatgaaat	gagtagagaa	160560
aaattcatat	ccgttggaag	agatgcacta	tatttgtacc	agaatgtgat	cacggatatg	160620
actggcccca	aacataacaa	gagactccgt	atccctcaac	aacaagctga	tttttgttac	160680
attatagcca	tgttggttaa	tgatgttccc	attacctcag	atttactttt	aactggaaag	160740
gcgacaaatt	tagtacaatt	tgettctgct	atggtggatc	ctgcatatcg	cctggctgtc	160800
cataaaatgg	cgtctgtttt	caatagtagt	tattccgtat	ataaagtcct	agatcttgac	160860
cataaaatgt	tattaagggc	taatttaatc	ctatctattt	tatcagctag	aaataagtgt	160920
ctcagtgaag	gaaaacctag	aacattaact	cagagcgtgt	atttgttctt	aaatcatctt	160980
ttgcggaaca	aattgaggtc	tagtggtttg	accagtgaag	agagttctct	aggaacagcc	161040
gtaaaattgg	tgtcacagca	attaatgtat	gaggggtgtg	ctcgtcaaac	catcgaggac	161100
ggatgtagca	tgaattagcg	aaacttttag	gatgaagacg	gtgtaacact	gaaatgtttg	161160
ggagccgatg	ttaaggacgt	gaagactggt	ggactatctg	cattgctatc	tgataggctt	161220
agaaaaaaca	ttagaagaaa	cgtccctttc	tactgaaaaa	atatggaaac	tactacgctc	161280
gctttcattg	caatgggttt	gcttttctgt	gtgttgatgg	tgtacatctt	tatacagtac	161340
acgtcatcaa	ataaaaacaa	gacagaggaa	agaaagggtg	gtaaattatc	atcatcatca	161400
actttcttta	cttcttggtc	aacatttaag	agagaagcgt	ctcctgaatt	aagagaagac	161460
ataccacttt	ctccacgacg	ttctggtcca	gaactgaatc	taactgacgc	tgaatttata	161520
gacctgaac	tcttttctga	agccagggtc	agtcgccgac	gttaacggaa	aaaagggtgca	161580
aaccggcaac	tcttctccct	cctttcagat	ataagtagag	gcgttgacc	gagccaagca	161640
tcattcgggc	tcattcaact	gtgctgacaa	caactaagaa	tatacattga	aataacccaa	161700
ataatcaact	tttcaagatg	tttggaagtt	ctgctaataa	ctttaatggt	gacaagaaat	161760
cttctcatc	atcatcagct	gccgcacatc	ctgacgatca	gcagttaggt	ccccttggac	161820
tgtctactgc	tgatttcaag	aaagtgtctg	ccattcttgc	caacagaaca	gagtccctat	161880
atctgttgcc	agattctccc	aactttaaga	atgtgattaa	caaccctaac	cagatatcta	161940
ttgtgccgtt	tctgggttca	tcgaaggcag	ctgaaagtgg	tagtgcaaac	aagaatgaaa	162000
accaggctga	aaattcttct	aaaggaggaa	gtgatggaaa	gaaatcatcg	cagcagaaca	162060
agtttaacct	attgaacaag	gtagaggctg	aagaaatggc	ctttaaacgt	gtggctgaac	162120
tcattgcgga	tactcctccc	tctaaagata	atccattgag	ggatgaccct	gatgctattc	162180
catcacgcaa	cccatgggtc	aaattgactc	agaagaatct	ggaatatctt	ttctgggaag	162240
cagtaacaat	tgaagtctcc	aatgatagga	gtatccgtag	tggaagatat	cttcaagcca	162300
gtgaagtggg	ggagaatcca	ttcctaatag	ccatcagtg	tgacattaga	atcctgcaaa	162360
gaatggcact	taatgtcgtg	tggttcttca	acagattctt	ccgcattggt	tctggactctg	162420
gagtagaaaa	cagggccaat	tccacctatg	tagcaactag	cgatgctatt	gccagatctt	162480
gggtagagat	gctcctcaag	aactttattt	ctggagaaaa	tgtgccccag	gcactgaagt	162540
atgtgaagga	acattatgag	catgtttata	acaagatcag	taaatgtgga	cgtcagccat	162600
cctattttgt	tgttgaattt	gaaagagtgg	acaacacgat	tggatttgtc	aattctgata	162660
ctgaacataa	tggatcatca	tacatggaat	acaggtgctt	tgacacaatc	aggaagaatg	162720
catcatctgg	gccaaagtga	ggcggaagaa	gtgggttcoct	gtcttctgga	acattcttca	162780
ttgataacga	aatggggaat	aacaatagta	gtgcagctgc	agcttctgcc	cctgctgttt	162840
ctgctggagt	ttctccatca	ctttctccat	ttagcagtg	tggagatgat	gatgacgacg	162900
attgtagtgg	tgatgacgtg	tgggggaaga	agatgatatt	caacacatca	ggagacggat	162960

caggagaatc	ttctgggcag	aacgggtggtg	gtgcatcaac	ttacaagaga	tttaggtgtg	163020
gagaaaatc	tgcttctctt	tctcagaagg	aaaatgttcg	tctcatggcg	atgccaaagg	163080
gaaatgaaga	taaacaacta	ctcaagaaca	ttatcaactt	tctgaatagt	gcacttaact	163140
ctgttgaaaa	ccatgtaatg	tgtacagatg	aaaatatctt	cgatgaggat	caagccgagc	163200
attatacgtc	aaacaaggaa	ctatacaagg	ctattgtttg	ctccaaccca	gccaatgtgt	163260
acagagtaat	ggttgaactg	tttgtcaacc	ttattcttcc	ccgtcttagg	aacccaattg	163320
tgagtgcac	tgaactgtta	caaaatcttc	cttcaaaca	tggatcggtg	agaacaaaga	163380
agatggttga	acatggatgt	acagatatgc	gctatgacat	ccctccatat	gcaaagggaa	163440
agattcgtct	tagtgccaag	agagcttgtg	agtgcagaaa	actgtgcaag	gacgtgaggt	163500
gctttgacaa	gtctagagag	gcaaattctca	cccctagcca	gaaggcagga	agggaggttg	163560
aggaaccttt	tccccgcaac	cacaattctc	acaggagtaa	cgctcacgac	tttactttct	163620
atgacaagta	cagggcaagg	atgaacaagc	tcaagaagga	ttcaaagaag	aaggtaaaga	163680
agattgacac	ctttacaaca	acagacgatt	ttctcctgca	agatagggaac	gcttttgatc	163740
tacttagaaa	gtgcttcctt	tctgcctctc	ttcatcacat	tttctgtcct	gatgtgctta	163800
tggtgcatag	aggagacagc	ttcaatatta	actttgcaaa	caacaaactc	gagtgtctata	163860
atgaacgtaa	cggaattgaa	gaagttactt	cctcacaac	ggtcaacgcc	aaggaagcac	163920
ttgaggatat	tacaaaaatt	aaaatgaaga	gaggggatga	tattatagat	gttgtgaaga	163980
gtaaaggact	ttctttgagg	gaattctcta	agaagggttag	taagattgtg	agaaggttta	164040
atgagatcac	aaaccaactc	tgcaacaact	gcaacgttaa	ctcttcta	ggagatgttg	164100
atttccacgt	ctttacttct	gtgtgtgtct	acatccacaa	cattattcct	gtgtctogaag	164160
atatctccat	ttttgcagaa	ttgggtgaag	aattgaccaa	gcttggttaag	gagtgtagag	164220
acgtggctgg	agaggacaag	acatatgatg	atattatccg	caattacgaa	attactgtaa	164280
agtactttaa	gctctttaat	gcactcgtaa	aattctgtca	caggaattat	aatgtggcag	164340
taacctctgc	cattaacagg	agagggtaca	tgtgcatggt	gagcaacctt	gtcgggtatt	164400
attgtaagct	gtctgataac	gctatccagt	atcacgaatc	actatgctct	ttgactcta	164460
gcactcttta	tcagacttat	tatacgtctc	gcaataacaa	ttctgaagat	ggaggaggaa	164520
actcttcttc	agaaaagagc	aatgcagatg	tagccaagac	tatggcctct	ttctatgacc	164580
agttcgataa	gagtgaagac	agcaagaaaa	ataagaacaa	aacttcaaat	gagatcctta	164640
taaaaatggt	ccaaatggat	agggtttttg	atggcatgga	tgatgatgat	gatgaagata	164700
gtgatagtag	tagcagttag	aatgaagagg	aggaggaaga	ggaggaaatt	gtaaagaaac	164760
cagcaaaaga	gaggaagtg	gaagatgttg	atagcaataa	gaagacactg	ccaaaggaac	164820
ctgccgttaa	gaaggtgaag	caggaagaag	atgtggagat	ggaggaaagt	aaggaagcag	164880
cagcagaaga	agaaaagaaa	gaggaacagg	aggcgaagga	ggaagacgct	actgagtatg	164940
acgacgatac	agaagaggac	gagaaagcag	tagcatctga	tgaagatgaa	gatgatgaag	165000
attctaaagc	tattttctaa	atatcatgtg	tataaaatgt	agtattttaa	aggtataaat	165060
aaacacataa	taaaagttaa	cattgttttt	tatttgacac	atgtacatca	tcaatttcac	165120
atggttatat	ttcttgtagc	ttgctcggca	gcaataagag	caagagaaga	gtcatataaa	165180
acagatgtct	gtgaagcagc	catgatcccc	ctcgatata	tattaataag	gggagttgct	165240
gctatatata	cagcataata	aatcaggttc	ttccaatccc	tagaagccca	cacattgggc	165300
atttttatat	aaaaatcaac	gtctactatt	gacctggcaa	cagccactat	acctagctca	165360
tcaatgttaa	caatcttccc	ctgaacaccc	attacaattt	ttgatttttc	tgaattctct	165420
cgtatagcag	caatacggct	agcttcatta	tttaaccatt	ccaattttac	acttgaaaca	165480
taagctgggg	atttttcatc	agttacttga	gatacaatat	tattccattg	agataaaaagt	165540
tcctccaatt	gagttgtagt	cttgctcgct	tgctcttgga	aactttgcat	gacgatagcc	165600
gaatctcttg	ttgcagtttc	gatgatggaa	atacagttga	gccattgcat	gtcagtaagg	165660
atatggcgcg	gattaaactg	aacaaacaga	ttctgtgcag	gagaaaagt	caagtcttct	165720
ggaatcttgt	tgaggttgat	attgtccttg	gtgtgtttcc	atgggatagg	tttctgtgcc	165780
tttgaagcaa	taaagtatga	tgggattcta	ctgtatttct	ttagtgtttc	tgaaatgttt	165840
gtctttatac	ggtcctttct	tccatcatcc	ctaattgta	tcaaaagttt	agatgtgaat	165900
agagaagagt	ctctaatttt	ttgcatttgt	tcattgtgaa	cctttatata	ctcgggaaca	165960
gtattgttat	tatttagcctg	cctctgttga	cgctgctgct	gttgtgtgtg	ttgagtagta	166020
cttccaatgg	tatttaattt	gtccaatatt	gaacgtaatt	ctacaggcac	ccttctacca	166080
gttccaagag	gccctccacc	gccgatggcc	gctgcagctt	cttcaactagc	ttcatcattt	166140
ataaaaagtc	tccccgaaga	agcatagcgt	aaagatccat	acccttcatt	tccagtatca	166200
atagtaatcc	cattttgttc	agctgttatc	ctttccagta	tttcttcaat	ggcctttcta	166260
ggatctagt	caacatataa	tccatcgacc	ataatatgag	acacgtgctc	gcacacttct	166320
ttaataattt	tcatctcttc	gttgctgaca	agcatcatca	ttctagatat	aatagtagca	166380
ataacattcc	taaattttgt	ctgaagaaat	ctgtcgatct	ctgtgaacgt	aggctcgaaa	166440
cgggcaggca	cattaactgg	aacaaccgct	tgtgtttctt	gttgtgtgtg	ctgtgtgtgt	166500
tggttgatcg	ctcttcttgt	tccaccacct	actatattat	tactgcctct	gagttggtcg	166560
gctgcaatag	caatgatgtt	ttgtacacct	ttaccagatt	ctgcaattac	ttccattttt	166620

WO 01/38351

47

PCT/US00/28888

gagtccttcgc	ctgaaattgg	tacagaagga	atcttggggtg	acaacctttc	tagagttgtg	166680
ggtaatccag	tatcatactg	aagaactttc	caaataaatt	cgaatccagt	tccctgccact	166740
gtgcccactt	tactggggaa	aatatcgatc	ctccctgtag	atagaggact	aacacattcc	166800
ttatagatcc	tcattgggct	taaatgttgt	gtcgatttag	tttttgaggc	ttcataaagc	166860
acagcagagc	agagaagggg	ttcaaagaat	acagtttcca	tgcccaattc	ctttaccgaa	166920
tatttgactt	gttttagatgc	atgaacagga	agatttgtgg	gctgagaagt	cgtaccaggg	166980
aggaatgcgc	taacatgttc	aattttccca	tcaaccgaaa	aacttggtccg	tacaccatca	167040
acgaatgaag	aagaagaaga	agaaagggat	ccagtgtcgt	cttgatgggt	attaatatca	167100
ccaatttcca	tatcctttat	accatcatcg	cgttcagtat	tctttacaac	tttaaagtac	167160
gacaccaatt	cacctttatt	tttcaatcca	aggaatccat	cttctgcgat	attctgacca	167220
tatgtcacat	tcttggttag	agatttaggt	ggtacataat	ttaaagaacga	ttttgaaatg	167280
gtttcagaag	aaaatggagc	agtggcagtt	gtttgtctgt	gttgtgtgtt	ttgttcctta	167340
gaggatatta	gaggtacatt	aatgaaaagt	tccagatcct	tattcatctt	taaaaaccgt	167400
cctgttggtg	caggaagatt	aactacccta	ataggggttt	ctgatgaata	tctggaggag	167460
gatgatggag	aagaagcgac	tgaagtggca	gactttaagt	tttcggctac	attatctcca	167520
ttcaaaatac	tctgaggtgt	atcggtcagt	gtatagtatc	tgagagacgt	atcattttct	167580
acaataaagg	gaggatctct	agggaaagca	aggcacacgc	ctccattatc	ggctgccgaa	167640
tgaataactg	cataatgtgt	aacaatcttt	gcaattagat	ccatcacggg	cttcttgaat	167700
aataaaaaatt	tgtgttgatc	aggatcggct	gtatatttta	gtagcacttc	acgtactata	167760
taatctgtta	gttgtctagc	ctcctttaca	gcccatacgc	ccgatttcaa	aacgcaatca	167820
aaaaatggaa	gaagaacaac	ttgaagtacc	tgctggtgca	tgtacatcca	agcggttcca	167880
tccactttca	ctggctgtcc	tctcataata	tttcttgtca	ttgtttctaa	acacgaccga	167940
taaagaccca	ccacgcagtt	aatgggatcg	taccatctgt	actgccgtcc	ataacttctt	168000
cttcggccca	gtataggctc	ccctgggaat	agtatggaat	cgcagttaga	ttctcttcca	168060
atattgtata	tggatgcaat	ctctcgagca	atatctccta	ccgtaatagc	agatccgccc	168120
gagaaaagc	ccacaacagc	atcagagaca	tttgagaatt	ccctcccttc	tagattttaca	168180
ttattataat	atatcagaga	tgctattata	gcagcagaag	attccccaga	atcgtctctc	168240
agaattaatt	ttaacgtagg	aatatcaaca	tctggtataa	aggctgtagt	attattttga	168300
ccgaatgtaa	gaaaccttcc	agaatcgtgt	atttccgaca	cgttactttt	tctgtcgaca	168360
acgagtgttt	tattatgttc	agcgtcgtta	agagagggtta	aaaaaagggtg	cattttctatc	168420
cagtcctgtt	tctccgatat	acagtcctatg	attgaagtgt	attgaggccc	aaagaagcaa	168480
gtaactgggt	ttttcatgga	aatggaaagta	ggtgcccctg	aagggacaat	gttggtatcg	168540
gtagaagttt	tttgccctctt	tgtacgtttt	ctatttttcat	cgtattcctt	agaaagggtg	168600
gaagatacat	ctcctaaaga	ttttccttta	gacgataaaa	tcgattgtgc	gattttttcc	168660
ttttgttcgg	tcgtcaatga	ggcggttgag	gaggaggagg	aggctgctgt	tgctgttgct	168720
gttggttcat	aaagtttggg	ttgaggttta	tttttaaatg	atagttgata	ataccgtgta	168780
accaaatcct	ttaaagaatt	ggtatttttt	ccaaaaaagc	actgggcccgt	aaagttagtg	168840
ataccaaatt	cagtcgagtt	aaattttgac	aactctggtta	tgataatgga	agaaactgac	168900
ttggagittg	aagccatgtc	atttaaatca	cgtagccttg	taagagaggg	atgttgatag	168960
agaattttct	ttgcattttg	gttcttactg	gtgtcaggaa	atgctatctg	gaggggggtg	169020
tttatccaca	ccttcacatc	cattagacga	aatatcacag	ggtgttgtaa	taatttccct	169080
gtgaatctag	aggtcgtaag	aacaatatca	agtatgccgt	gatgggcata	aaacacacca	169140
taggcgtctt	tcaatactat	tgaggaagaa	agtcctttcca	tttcccatc	ttcttgattc	169200
agatgggata	tattgacaag	ttcttccatg	cogtcgtttc	ctttaagatc	tagccgtatt	169260
tctttctcat	catgcattgt	ccagtcctta	atattaatga	caacaggcaa	aagccctgcc	169320
aatctctcta	aagacgaagc	agtttcttcc	atgttcagct	tgatgttgtt	aatttcgcct	169380
tcgataaacc	tctcattttt	acgagaataa	catectcttc	ctcctcgagg	aaatttttgt	169440
tctgtcagag	atgaaaatcg	tcgtgcttct	cgtgctcttt	tttccactaa	tgcatthtaga	169500
gattttgttag	atattttcttc	aatagtcate	tctttattcg	tctctataat	aaacacagac	169560
gggtgtattt	ttacccttat	ttttgcatth	tctgtttcac	ttcctaaacg	cccctcggtg	169620
aggtttgcaa	gggtagaatc	ataacacgca	aaagttcccg	cattaacaag	cggtagagaa	169680
atttcgctat	cgccatcttg	ttgttgttgt	tggtgtgtga	gcccttcagc	caataggcaa	169740
gctaaacggc	tgtgtctctc	gttggtctta	ccaggcaggg	gccgttctat	tggttgcggt	169800
tgcggtgttg	acataatttc	ctcctcgacg	tttctcgtea	gcctgtaatt	cgagtagagt	169860
ttgtttcctt	atacgggtgt	gaaccaatag	tcggtaaaac	accaatccac	cggttacgat	169920
gatggatata	aacaatatat	agaatagggc	cttgtcacta	gccttgacgc	attgtctctt	169980
tctgcagctg	cagtcaaaatc	tttgtagtct	gggcctaatt	cgctctgtac	gtgggtccgt	170040
attctctcag	acacgtccat	gttgttggtg	gataaaagta	taaatactca	ggtgtacata	170100
aaaacctctc	acatgatggc	aggaatcggt	agaagagata	atagaccctg	attacatctg	170160
gacattgatc	caaataagga	aataccctat	aacgttccac	caacccttat	tatctgtgaa	170220
aaaaaccctt	tcgtgtttaa	catgcagaag	tgctcagact	gtgctccttt	ccctccctac	170280

cccgccactg	agaagccttt	ccctccatac	cctgggtactg	cagtagaaga	ggaggagaag	170340
caaaaggaaa	ttgaggagct	tctgggtgac	caatctttcc	ctccccatt	ccctggaaat	170400
aagctgagag	atatccccag	aacctaccct	ctcgaatttc	ccgagaagaa	ggagaaggat	170460
ttcccttgcg	ttgacactac	cggtcacagc	gatatccctt	tcatcgatct	ggagaaaacc	170520
ccaccccgta	gtgacgttag	gcacgggtac	cactacttaa	tcaaccccaa	caaggttggtg	170580
gagcttaacc	atatcggttg	taagctcact	gaaaagcaag	aaaacctgaa	caaattggtg	170640
ttggatgttg	atgacgttgt	gattaatctg	tcaagcactt	tgaaggaact	tgagaagctg	170700
cgagctggcc	tgtgcaagtt	ctcaaaaaac	tagacggtaa	aggtatgacc	aaaccgggtat	170760
aaccaacttt	ttatataagc	ctaataattat	aaatcaacaa	ctaacaaacg	cgaaaactaa	170820
caataggaac	catggcctca	tcatcatcat	ccctgtttgc	cctctcctct	gtcgcacatc	170880
ccgtgatgat	ggaaagagat	gaagaaaata	cactgtccct	caggaaacagg	aatgtgaaca	170940
aaccaacacc	tgttagcgcc	gcctgggtgc	ctgttgatga	agaggatgaa	gatagggagg	171000
aaatgagaag	acttgaagat	ttttcttcag	atgaagaaga	cgatgataat	aaatcatgtc	171060
attgtgacca	tagcgatgac	gatgacgatg	acgaggagga	tccttcatgc	tttaagggtt	171120
tttcagctgg	cctgtgctct	tttgtgaggg	gtttcttttg	cttcctcagg	aagtcactta	171180
ccaagaaaca	ggtgttcctt	cttacaagcg	cagccgttgc	tgctattttt	aagactagag	171240
atgtggctaa	aactgaagaa	ggcgcagcaa	ccatggaaga	aaattcaaca	gatgtgatta	171300
ctggaggaga	tggagatagt	ggtattgctg	ctgatgttgt	ctctctcgct	agtgaggagg	171360
agggagaaaa	tggatctctt	ttggaatcta	ttgcaacaac	actcatcaag	actacaattg	171420
aaaatcttgt	agatgggtga	gaagaaacca	cagaattgta	aaatgttttt	tggtatacaa	171480
ataaaaagaat	acataattta	tataatattt	tattgtttcc	tcttaaatga	caatttcctt	171540
atgatacttt	taggtgaagg	aataaaggta	ggagaagggtg	tggaagacac	aattgtagtg	171600
tttccataca	catgattctc	ccacatgtta	ttattgttag	gtgggaattt	cttaaacgta	171660
acttgggtta	tacgtggagc	attgttggtg	ttgttactac	tgctcctggg	tagggtgttt	171720
ttcttgtgtt	tcatctcctt	caattccttg	atacacttct	ggatagatgg	agattgttga	171780
ggtacttctg	gcacgtattc	gtagacatgt	tcttccacat	tatcatcctc	tccctcctca	171840
ttcttgatag	tgaagatggg	agaagggtggg	tttgaaggaa	gcttcagatg	tcgtccgttt	171900
acatatgaca	cgtaatcacc	gccattgctt	ctaggtggag	atggcggcgg	tgttggcaga	171960
ggaagactat	tgtcagaaga	aaatgacgac	gagcagctcg	acattggcat	atagacactt	172020
tcttcctctt	tatttgtact	tgttggttgt	tgttgagggtg	tattggcaac	tggcctgggt	172080
gtaggcggag	tagcataggt	tggaccagaa	aatgttacca	gttttgtgac	ttgaggggta	172140
acattggcca	gatcgtaagt	ggtagatttg	agcatgggtt	cagagatgtc	caggttcctta	172200
tctggagtgt	tgaaccggac	gtggctgatg	aaactggatt	cggtgttgct	cccgttcctg	172260
ttcttcttat	tcttattctt	ctcccacggt	gagcaagaaa	tgcaagtagt	agttctgtcg	172320
atgcacttca	cgagcatggt	catggccagt	gcgagcccaa	tagtcatgga	gcagccgact	172380
gccaatccac	tcaggaaaagt	gacgaggata	atggtgtcgt	tggtgataag	attgtccatc	172440
tcgagagatg	tttcttcctc	gactaccaga	cagatatatga	tgctctgggg	ccctgcttcg	172500
tcctttttata	ccgtcacttg	gttagaaaaa	acgatagtgt	agttcacagt	atttgtccct	172560
tctggtcttg	ggattggggg	tctgatactc	tgcccagcct	ccagatatag	gtggatagac	172620
cgagattctt	cccttgatgc	agtttctatt	ccatccaaga	tatcttcacc	aaataagtcc	172680
ttatgcaagg	acaccagaat	attgaactcg	tgaacgcttg	acttgttgat	gatgcataat	172740
tcaccttggt	cgactgttat	acataactct	tcaatagaat	cgatatagcc	ggggtttattg	172800
gcacggctct	cttcaggtaa	gaaaactgca	aatattgtat	taggcacaac	atctggcacg	172860
tttacagaat	gacagtaatg	tttatctata	atgtcaccag	aaaagtgaac	atctatcatc	172920
cgacactgga	gaaagggttg	aggaaaaatc	ttccgcatat	tcttgagtgc	ccgcaataac	172980
ttcttggttc	tggagtttga	cttttttagga	ttatttggtg	ccagacttac	agcatcgaaa	173040
atactttctc	ccattctctc	agcgggattt	tccatcttac	ttttgtagtg	gttttcagtt	173100
cattaattgt	ttttcaggga	ggctataaaa	caaaactctt	acgcatctgt	tcatttcagta	173160
gtggaccaga	cgtcgagatg	acgagacatg	gtgtgcttgt	tccaaaaggc	cgttctaggc	173220
atgttatcct	aggaaatgtc	gactacactt	tctgtactac	tgacaataat	tgtgtcagtc	173280
tagacatcga	tttcaaggac	aatatcacag	acaaaaacat	tcagttattg	aacaagaaat	173340
tggtgaagaa	aacagcaaag	aaaataaaga	aggaagatgc	acctgaaaca	aaggaaaaata	173400
gtgacgaaga	catatacgcc	accaagggaat	tcgaacagac	aataaaaagg	ctacagacaa	173460
aaaaagggtg	caccgaggaa	aacgccatcg	cggccgcagc	tgccgctgcc	actgctgctg	173520
cggtagaaaa	ggctatgcta	tcagaaaagt	aaggaaaatc	aatggtcac	aacagagcta	173580
gaatgggtgt	ttctaagcga	gacacgtccc	agaaacagtt	cactgcattg	aagaacaggg	173640
aatctttctt	cagtgttttg	atatttgaaa	ctggatcagt	gatagttgtc	gggcttcaag	173700
atccttctgt	tacaaaattg	tgtgtgatta	aagccacgac	tgatattgct	gatattctac	173760
agaaaaacat	cagtgtggct	aacgtgtcta	tagtgaatac	agtgtccact	tttaatatag	173820
tccacttgaa	ctttatttcga	ctcgggaaat	tcttcgaaaag	aaattgcac	tcttacagtt	173880
ataaccacga	aacgttcccc	ggtatgtttt	tcaagctgcg	agtgcccgca	aagcctctct	173940

tgccctggaga	gactataggg	gaatactaca	caaagggttgc	aatgatgcgc	gatagtaagg	174000
atcccaattt	taaaatgtct	gactgggttga	ggataaaaac	tgcattaaca	tttaaagttg	174060
ggaaaattac	tgtgctcggg	gaaggagaga	gtgggtgcg	tgatgtttct	gtcgtatcca	174120
aattactatt	tggtttatcc	cattacttta	tggacaacaa	cattaaaatg	tcccccaaa	174180
aagcacaaag	agtcagagaa	aaatacggca	tcccgcattc	agaatggtac	ttgtacattg	174240
acatgttgct	ccactcctac	ccgtacgtca	aaccatcggc	cgagcaagt	aaaaggcgca	174300
tgggtggacca	acaacatatt	tctgaagtgg	ataggacata	ctatggaaca	aagaacagta	174360
tggacgctgc	catgtctgca	aatttagtgc	cttcaaaaaga	agagagtatc	tccttcatta	174420
aaaaaataag	atcacacaac	ctttttggac	atttgtgtaa	accttcaaaa	gaaactactc	174480
gacgtgctat	agacacactt	tcttttcgatc	ctataaacca	agacaggtgg	tggaataaaa	174540
atgaccaata	ttacggtaaa	gagagatgtg	acccgttttc	tgttgacagt	ttagtgtctg	174600
tttctgaaaa	tacaaacagt	atgatgaata	gtcgcatttc	ttgtcaagg	aaatggtggc	174660
tagatgaaaa	cgaatacaag	gataaacttg	atcatattgt	ggatttgtgt	acagaagaaa	174720
tagtgaggga	atgtgaatca	aagggtttta	ttgcctcccc	atttttgagg	aagcaccaga	174780
agggaaaaat	accaacgcct	tatgttttat	tagcgagagc	ctgtaatcaa	aaaaatggta	174840
acaaaatgag	tattaaacat	aatagtaact	atttgtcggg	ttcaagtagg	gcgaagagga	174900
atgcaaaact	acaggaaaaa	caccgtgtaa	ctttagccag	gttgaacacg	atgatggcgt	174960
cgtaccgatt	tttgaataat	tacatctcaa	cagacattgc	acctgatttt	gccaaagtat	175020
ttggtaatga	tgtatatagt	ttattacatt	taatgacaaa	cttgcctaaa	tcccgtggac	175080
atgctctaac	atacaacgaa	agggcccttt	caagtaatga	aagtacatat	aaaacacctg	175140
gaaatgcata	cttttagtact	ctatttgaaa	ataccattat	aaataaccaa	gaaactgcta	175200
ataaaggtaa	caatagaaaa	cgtaaatttt	ctcgaaatcg	acaagaaaag	agctcttttc	175260
tgtgcaacgc	gtgtggtgtc	aatttgaaca	agggtagtga	tgaaatcata	aagggtattt	175320
gtacaagttg	cgatcaaaat	agtaccagtt	acatagagaa	tgcattatct	gacattaaca	175380
gagacaagaa	gattaaacgt	tttaaagcag	ctgcaaccca	tccgccagtg	aagcaagaat	175440
tggttagattc	tttatcctcc	tcttcatctc	cttcttcttc	ttcttctcag	acgtctaaca	175500
agaacaatag	atgcacccct	agtqatttta	tagattatgt	gtacaaattc	actgacgaaa	175560
caacaggtgc	tccaaagggtg	ggcttagtgt	ttaaaatgtg	tgatattctt	gcacccctag	175620
caagcaggag	agggatggaa	gatcgtccca	cagccaacta	tagaacctcc	ttacattcag	175680
ctactcaaaa	taaaaccaat	ttgaataaac	tattagtttc	tgctatcaag	gaaacaggag	175740
ccactgaaac	tgaagcacag	atattcaaca	agattattgg	tagtgaaaag	ggactatcaa	175800
ttctctgtca	acttgtggaa	aggaggaaca	aagacaataa	tgtcttcgac	tgatttgtct	175860
aagaatgcct	tccatgactg	gggtggtatca	aagacagatt	gtgaggtgtt	tgatgtacac	175920
tgtgagacgg	acagagattg	tggcgctgct	tgcgagaaca	cgtactctgt	tgacggaaag	175980
gaggttacaa	aattctcttg	taaccaacag	tcgggaagat	gtgccaggag	tgtttatagt	176040
gcgtcttctc	tagaaagagc	agccaatgat	cttggccaca	ttataggtat	catcaagaaa	176100
aatccaaaat	tggaggaaga	actccctgaa	ggtttttgtt	ggtttatcaa	tcacaattga	176160
ggagatttgt	ttgtgaataa	gcgagccg	tactacgaca	cgatgcattc	aagcataggg	176220
aaactggata	atgtggacac	tcttgcccag	ggtttagata	aacggatggc	ctcatcattg	176280
agagagcacc	tactgaggaa	gttggaactc	atacttttac	aaattgataa	agttaaatat	176340
gaaaaggcaa	agaaatggat	attggatata	acacaggagg	ctggcaccga	agaggacaat	176400
aaagaagaag	aagatgcgaa	aaaagaggat	caatctctta	gcgttagtga	aattgtggat	176460
gttttaacgg	gcacacatga	ccctatgccc	ctgagggcta	gagggtttat	ccagaaaaaa	176520
atatatcctt	tgtcaagaaa	cgaattgaga	gaattagccc	ttaagggaact	tttccctgaa	176580
gaaactacat	ctcctcagg	tttgagtagg	caacatgatg	tatctacgcg	tgaagattta	176640
tgcaatgaaa	gtatgaatgc	agggagggca	gaatccattt	ttagcgaccc	tgattctgga	176700
gagtagctgg	ctacttgtgc	atgtctttac	tcggaaatatt	taacagggcc	tgctgttaag	176760
cacaaaacat	acagggtatgt	tatagactac	gacaaaatgga	aaaggactgg	aagacctgaa	176820
tttctaactg	atcctgtact	tcatttttaa	aaggcagaag	ctgtgtgtaa	atcgacaaat	176880
ccaaacttga	gggcaattta	tagtccagat	aataaagggt	tcttgtgtgc	gcctgtagct	176940
gaacttgtaa	agacggcatt	aacttttagg	ggttcacacg	aaccgtctct	cattgtcag	177000
agagatatata	atcaagctga	aaatctacca	tccaattcat	ttggtgtaaa	ctggccctat	177060
tggaatctcc	taaatcgcat	tcaagaccag	tacacgtta	ttgataaaaa	aagatcgggg	177120
aaaatgtctg	catctttaat	attggacgaa	tacctcaaga	agactgcttc	agccgttctg	177180
gatgtagctg	actcatttga	gaaaatcaag	ggagaaatcc	aatcacctga	ggaggctgcg	177240
gctctttctg	ttgctctcta	tggagcacct	ccaaaacctt	cagcttcggc	tgtggcctct	177300
atcatcactg	gagaaagaac	atcttttaaac	gacaaatata	tatcgataaa	tgtcctattg	177360
aaaatgtctg	ttgtcgcgt	tggacaagaa	aataatcgca	agagagccga	ccaggcagct	177420
gatgaaatta	gaaccatcat	ggaagatatt	tgccgggtgc	tgccgggtgc	gtacaggcaa	177480
tatagcccg	tcgaggaaga	aaataagggtg	catataggca	tcatgaataa	caaaacgcct	177540
agcattgttt	gtggatatta	tacaatggac	acatctattt	cttccgaacc	tctttctcta	177600

acagatttttc	aaaacccccac	tgtcatttgcc	aatgtgacta	agcggatgga	gagcattttt	177660
tcaaaggtcg	actctgctag	gtctacaaga	ttcgacgctt	ttgttaatgg	tggtgcaaat	177720
aatatggata	taaagtcac	aatagattgg	gcaaataatg	tagaaaatgt	gatcaaatta	177780
ccagatttcta	cacctaaccc	ttgttcagtt	gacactattg	tgtccagaga	cgcaagtgtg	177840
gttaaaaacag	cagttaatga	tatatacgct	tctgttgga	aatcttattg	tcgctctgca	177900
acacagctaa	cctttatgag	cgagattgaa	aaactgcgaa	aggctgcagt	tgtatgtttt	177960
gaggcactca	tgtccgatac	tagggagagg	gcattcgtag	agttcctatt	ttacgttagc	178020
tttaaggaag	atgcatcaaa	taccaattca	aaattgtttg	ttcagaataa	gctatcttcc	178080
atgtctggaa	accccagaca	gcccataaaa	ttggtacgcc	gttctgctga	ggaaacacta	178140
ttcgggctct	gtttcatggt	taaggtaatg	cctccagaat	tcatgaactg	tatatttaac	178200
ttccctacca	ttcccccattc	aacacaatac	catgggtctat	atggtacatg	tttaacccct	178260
ctacttagaa	aatacggttc	ttcattcgaa	aagtcctggg	ctcattttga	ggaaatttta	178320
agcgaaagag	ccaatgcagt	gaaaaaattt	ggtgtaaacg	atacagggat	agattgtcta	178380
gatgcagtag	caaatctcac	cggacctgtg	tatgttctca	ttttagatct	tgtacgtact	178440
ctaagtgcgc	agagatcggt	ttcaactaaa	tttctccgtg	aaattaagga	aaactatctt	178500
ttgtggaata	ggtttgtgtc	ataataaaaa	tggctcagac	atcaaagatg	ggaactaaca	178560
agaggtgttt	tgaggaggaa	gtggagggaag	aaaggcaaca	acctttcaca	aagaaatcta	178620
aatcggaacc	accagtttt	gaagacaaga	gttcatccac	atcttctaag	aagaagagca	178680
aatccaataa	acacaccaag	accaagggaag	aacaacttct	agaattcgtg	aaggatctgg	178740
agcggagcga	ccccactggt	cctgatgaga	aggtaagca	agaagtgtga	gaaaagtccc	178800
ctgaagctat	tgtgaaatt	ttttcaatgt	ttgggtcgc	tcaagacagc	aagttcaaga	178860
gccttcttcc	cattgaacgc	ataaagagca	tcactactaa	aattgttatc	gatgcaatta	178920
atcagcctgt	gcgcaagatg	ttggttgacc	acctctatca	ttttaaggag	atgcagaatg	178980
ttgtggagaa	atataaggac	gatagcgacg	aaaaactgag	cgtcattctt	aagagtaaga	179040
aatcccccaa	agaatttgac	ctctcctttt	ccgattacgt	tgatcgccct	aacaggattc	179100
tggttggtgt	aatataagag	gtggccggag	ctattgaaag	taagggaattg	ttgcagagta	179160
acagcatgat	catgaacagt	gttctgggtg	ctgtttgtgc	caacattcct	tacaacatga	179220
agattaatat	ttgtgtgttt	ttgactaact	ttatttgtac	atttgctaata	gatgatattgt	179280
acacattctt	tagggatgat	gagaaatttg	taatgagtc	ggtaacaaga	tacatttcaa	179340
aggattagaa	aaataaaaaga	tggtataaaa	ttacttgtat	ttttattcaa	acaaaacctt	179400
tatatattgta	caattccccc	tttacttctt	cttgatttcg	tccttgatat	cgatcacatt	179460
ctttaggtgtg	aatccaatgt	tcatttgtagc	ttcgaaggct	gcgcgctctt	caaatttgct	179520
gcacacgtca	atgagggtaa	tttttgaaga	tgaagattta	atgtccttgc	tcgagaagta	179580
agtgttcgac	atagtatttc	ccttaatgtt	tcctgcaata	acagttggcc	gaacagtaat	179640
gtcgttgatc	aagagatcga	cgccagtgtt	gttgaggatc	ttcaaagtca	tgtagaggtt	179700
ctcggcccca	gtgccagcgg	gatcagtcct	tggggctgtg	acggtagaga	tgacaagatc	179760
agcagaaaaca	tccttcattgt	ccttatcgga	gagaccattc	ttcatggcca	cctttccaag	179820
aggagtattg	taggacctct	ctccacgaat	tgacattacc	tttgcccttc	tttgaattgg	179880
gactcgcac	atctgatcat	aattagcgac	gacgcttctt	ccaacacgtg	tggtgaatac	179940
aatcattata	accatgataa	cgatttagagc	aatgattgca	atggacaaga	ttgcaataat	180000
tgcaacgtcc	aggtttgtta	ggttgccaaa	ttccattttt	ctttgtttta	gatggaagtt	180060
cttcccttac	gtgttggttg	gatccageta	gttagttctt	ctgtagcacc	atatacccg	180120
aaagggtttg	tttttaaaat	ttaaccctac	atcgaaccgg	atgaagtgtg	ccttgccgtg	180180
atctaatttc	agtaacgata	cctcccacta	cccctacttt	catctccatc	tctacccctt	180240
ccctactcct	cccttcacct	cctcctctcc	ccctttactc	gtattgggtc	tcttcacgtg	180300
cgtcttctga	attggatgaa	gaattgtaat	attttgtgtg	cttcattctt	ttcacttgct	180360
ttgctagaga	attgacgcta	gatttggtat	ctaaaaggtc	cataaaagtc	ctccgtctag	180420
aagggtgga	agattccttg	tcaatataca	cactgctagc	gctactcatg	ttggaattga	180480
tgtcatttga	aaaagaaaac	gttcttttcc	taccaacatg	agaagataca	ctgctaggca	180540
cgtccattgc	tgttgctgct	gctgggggtg	agggagaaac	atcttgattt	ttagaaggac	180600
ggtacattaa	gtttcgcata	ggcacaacag	atgtagattt	tcggtgctta	ccctttttat	180660
tgtatttctc	agatgcttga	tttagcttat	attttatgca	gtacacttca	atggacaatt	180720
tagcttttct	taaaagtttg	atcactactg	cgcttagcaa	gtgtttcctc	ttgtccatga	180780
aattgactag	tgtatttaat	gcattatcaa	aaatagggga	tctgtgttcg	atgattgatg	180840
agcctgcact	gtatgctaca	ccaaaccttc	ccccgccttg	tgtctggaga	tcaaacatac	180900
tattcttgct	gttattgtta	ctcactccag	aatacgcgtc	accgaaaaag	tggtttaatg	180960
attgacggtc	attactgtcc	ccgtatctta	tattatcgta	tactcgacga	attgcttccct	181020
tcttgctcct	ctctgaaaaac	atggtgccgt	cttcaactct	tttctctgcc	tcacttccat	181080
taccatcttc	atcctcttca	tcgactgcct	tatcaaattg	tatactgttt	aagacgtaaa	181140
tgggagggac	agccatttta	gaagaacaac	tagaaatact	atccaaaaga	agaactttat	181200
aatcttttga	caaagtgaac	tcttttggcc	aatttaatag	tacaattggg	atgttcaaga	181260

aaattggttaa	ccttcccacg	tccatacata	gtttgttgat	gtgtttctct	atgattgtag	181320
aaattgcccc	gaaacaaatt	tctgccccta	tagaacctaa	ccccatttct	tccattatgt	181380
ttgcagattc	tttactcatc	tgtttttagtt	tagatgaggt	agacaaatct	gatacatcat	181440
ttgtttcagt	ttgatgtttt	tcaaaatcct	cttggtgctt	gttgggtgtc	ccgaacagtg	181500
ccactgatgc	cttggtcact	atgtccgaaa	gcataattgac	taatctagac	caagaaacag	181560
cacttgatgt	tttaaaggat	atgaggttgt	ttaggaacgc	cagtctttga	ccaagcttat	181620
acgtttgaca	tggttgatta	tcgtcgtctt	ctttcagtat	atttatcagt	tccatgtgca	181680
aatttgcttg	agtttgatta	ctataaaatg	aagaatgctc	caccacctct	cctcctcctc	181740
ctcctcctcc	tctcctcctc	ccagcaatag	actcgtcgcc	ttcagcacat	actccaccag	181800
cagatgaaga	catggaagtt	tcacctctgt	ttaattcaga	tatactttca	acctttttaa	181860
gagaaggagc	aggagtcgat	ggtcttattt	tgctagctct	ttccccatt	atataagttg	181920
acgcgtatgc	tgctgaaaaa	agaggggttg	acatgatttg	cttctcttga	aacaactcca	181980
atttctcatt	tgttttcctg	aacagctcgg	cagattttaat	gtctgtcacc	gcagcagaat	182040
ctaaaaatgc	cttcatgggt	aattctgcag	cgatgatatt	tttaaggcgc	tccattactt	182100
tctcggatgg	ttttactggg	ttttctggat	cttggaaacac	aactttcacc	agttcagagc	182160
attgaagggt	ctttgttggt	tctagaagag	tttttaccct	tgtttttaag	attatagact	182220
ccatagtctt	gtgtaatctt	tcttcggcct	ctttcccacc	taatgtgttg	ttaacctttc	182280
caaccactgt	actcaaaaag	gatggattgt	tggtgttggt	tgaaaaataa	gactgatttg	182340
taggtggagg	tgagggggga	ggaagaagaa	tattggtagg	tttgagttga	ggagggtgat	182400
gaggtgggtg	tgagggtgga	ggaggaggag	gaggcatcag	catcgaatct	tggtccatga	182460
tgtctctact	ttcttctccc	atttcattgt	ctggaggaaa	agtcacgata	gatggacag	182520
gcgggtggtc	tatatcgttt	agttggtcat	attgttctct	acctccacta	ccaagaatag	182580
aaaacattct	aggtgctttt	ggtgaaggtt	cctttatctg	ggggagaata	tctgtctctt	182640
tcatgattgg	agggaaacaca	actgcctcct	cgcactcttg	tatatctgga	gggggagatc	182700
ttcgaacctc	gggcgcattg	ttaaattgat	gcatagtact	tattttcttca	tctcagagt	182760
tgtgccgaga	ttcatcttca	tattcttcat	attcttctct	ttcttcaata	tattcctgtg	182820
gagattgaac	agactcttta	gtttctacat	gctcttcaga	tttttggtga	ggttctcctc	182880
cctcctcttc	ctctccctct	tcttcatcct	cttcttcate	ctcttcatta	tcattgccgc	182940
cacctgctat	tgaagagaag	aaataatcca	ttccatcctc	attacctcct	tcactcatag	183000
ggctggattc	aactctatga	gagtgcacgc	tttcattttc	tattgcatct	tcctccttct	183060
cgggttcatt	tttttcttgt	tctataggtt	ctttttcttc	ctcttctctc	tctgttctt	183120
cttctctctc	ctcttctctc	tctgttctt	ctcttctctc	atcctcttcc	tcttctacgg	183180
gttctatttg	ttggttatca	tcttctttgt	tcttctgttc	ttcacctgct	gatccgcccc	183240
tggcgttcac	attttcaagt	aaaaagttca	acatgggaag	aacttgccct	tttgccctca	183300
ctggcgaggg	ggtgttcaat	atacttttcc	actgggtttc	aggtataatt	gtcgcattct	183360
cttggttgatc	cgctagcgcc	aataacgcag	ggtcagcctc	atcttcttcc	ccttcacctt	183420
cattgtcctc	ttcatattca	tctccttctt	cttctctgtc	ttctagattt	ccgtcatatt	183480
cacccctctc	ttctcctctc	tcctcttctc	cttctctctc	ttctcttctc	tcttcatcat	183540
catcatcgta	catgtcatta	gtgacatctc	tcacaaattt	actttgtttt	ttagaagatg	183600
gaggatcttt	atttcttcca	tttctaactc	ccgtgtctct	atcttcatct	tcgtaactct	183660
cctcctctct	ctcctctctc	tctcgttcca	tttctctctc	atcttctctg	tcttcataca	183720
agttactata	atcaacatta	ttattgatat	cgtctctgtt	gtcgatatcc	aacactcccg	183780
caccttcagt	tatacgcaga	tttgatgccg	cttccatgat	gaaatttttag	tatacaatat	183840
tgtgttaacc	gtggaaaatc	aaacactaca	tttactaagt	tcctttctctg	aaaaaacgag	183900
ttacctctat	agggtaatcc	tcagggtgta	tattctatag	tttggtcttg	aacgtgcttt	183960
tggtttttgt	tataactttc	tctaaaatct	actaacggca	agagcattga	cagttgagtg	184020
ggaaggggtg	gttcagagta	agacccaaaa	taaaatat	tttgattgtg	tgaatttcga	184080
acggtttata	ttactatttt	agggtactag	taaaacgttg	cagaaccaat	catgagggtt	184140
atcccgatgt	tgataattgc	gcttatcgtc	gcatttggtt	ttgcagctct	tcttactgcc	184200
aattcaaatt	acctggatca	caatatcaat	aaggaaactaa	atttgaccgc	ttcccttcaa	184260
cttcgggggaa	cattcacacc	agaagatata	gctcataaca	acagaattct	cccttccaag	184320
ctgagtgttt	tagaaagagg	gtcaatcatt	ttagcagaaa	tggaacaagta	caaaaaatga	184380
caacctacag	taaaataatc	tcaagagagg	aggaacattt	ctactcctca	acaacaacaa	184440
cagaccaccc	cctcatccca	atcatcttcc	caagttgaac	tttaaattct	ataagaaaat	184500
gctgatttat	gttctgtgat	gtgacgagtc	tcgcaccgac	gagaaagaag	attttaacaa	184560
cgatgaggaa	gaaaaatat	taaaggagga	agaagttagt	ggaggagaaa	aggttgtaaa	184620
tatcgtagta	gggagtatta	agatgatgaa	acctggttca	aaaaatcaaaa	caaagtacag	184680
gctctactcc	ctccttgcaa	gtgcgatagg	gagtagaaaag	aaagctgaag	aatatattga	184740
acgtctatac	aattcctttt	cttccatgac	gttgaatgaa	ggaggacgct	tggttccagc	184800
gatttattgt	ccatcctaca	acaataaaaag	gataaaaaaac	aatcgttcta	ggcctgtgaa	184860
gctgatccat	gcctccaggg	aattgtttatc	tgaaactacg	gtgaggggag	aaatactacg	184920

gaaatctccc	gcttcttctt	cttcttcttc	aacgtcatca	tccacttctt	cgttttcttc	184980
aatttttttg	tttgtagcct	ctaattgtac	ttctaaaact	gtatgcgatt	ttgtaaaaca	185040
catcacgtag	gaagaagata	ttaaccgttt	aagggtataac	ataatacata	tatctgaaga	185100
aaactatgct	tcaagatttt	ctaaaattaa	tagattaact	acatgtatac	aagggtataag	185160
taagacaata	aagattctat	aaaaatcaac	atttatttgt	atcattttata	aaattccata	185220
tttttgtagc	catttcgtag	tcgacgggta	gagatgggtc	gtagaaagaa	gatgctgctc	185280
cttttaaaag	gttattggac	atggctatca	catacatttt	cttattgttt	tccttattga	185340
tgatatttgt	catgacgttg	ggagtggcat	tttttggtgc	tggtgtatga	caaacggcaa	185400
cattactgtt	gttcttaacc	acctgagcaa	acatggctgt	aggggattat	ctttctatgt	185460
cctctgtagg	ggaagcaacg	ctcgtgggtt	ttatgatttt	aaattttata	aactttgtca	185520
caatattatc	tttaataata	tatgccgtga	cagacgttta	tagaagatgc	aagcggccgt	185580
ctaccaatgg	atattctggg	tgtacaacca	atgtgggttc	ttcgacttta	caagaagcta	185640
atctcggtac	cactgaaaaa	gataaacctg	ttcaatttgt	tagaggttta	gtccccagaa	185700
aaatgatgga	aaaatataga	tcggacttgt	ctcctaagaa	cgtaggggaa	tatattttac	185760
cttcagaaaa	agaaacagac	aaattgaaaa	gtgattataa	aaagggtaag	aaggttggtc	185820
ttttaactgc	cctgagtaat	ggcatgaca	gtaacaagag	gattataggg	ccaagggatc	185880
tgattagtag	agatgatgtg	aaggacaaaa	gttatgtctt	taagagattg	agcaaagatc	185940
cgctcgtcta	ctactcttct	gcaacctcta	aatacgtag	aaaattttcc	cctttcagag	186000
caaaaaaatt	catgacatca	acacagttgg	ggagtaagct	cggtgatcct	caccctatac	186060
ggtaggttac	tgtttttgta	ctaccacagg	gatacgtgat	caacaaagca	tacggaatgg	186120
ataatgagga	tctacacact	tggaaaccca	ctcttctctc	tgtgctcgtc	ccagactcta	186180
ataatgatag	attaacagta	gaatgtgcta	aaacggaccc	aacacataga	atcggcatct	186240
acggcttttg	agggctctgat	gataatagac	gtgcaaaaaga	agaaggttat	gtagaaatgt	186300
tgttatgtaa	ttgtgacaac	cacaaggact	tgcttaaggc	tcctctaatt	acagagtatt	186360
ctacaaatcc	aactgaaatt	caagtagatg	ttgctgcaaa	acgtgtttta	ttccctgccc	186420
ctggttccga	gcctgtaaaa	tcttcccaag	tgacatctgc	tgctcatcaa	ctagacggag	186480
ctactggcga	gcacgatatt	tcccatgagc	ccgtgaagct	atcagatacg	ggtgactatg	186540
cagttggatc	acccattgta	ttcaagccag	tttatggtag	atcttttagta	aatcttccag	186600
aaacaggatc	tcctctggca	ttgaactgcc	cctgcaccga	caaggctgat	ggaatatatc	186660
aagtcaatca	aaagggaggg	atattatata	gagatatggg	ggggatatct	aacgccaacc	186720
ctgtggaagc	tgcatacctt	tcttccctcg	actcttcttc	gtgggtgaca	actggtaaca	186780
aaatatcttc	ttgtacatgt	gaaggagaaa	aaataaagaa	aattgtgtaa	tcatatgtgt	186840
agttttatta	ccactaaatt	tattggtctt	cttggtactg	tttactactg	ggaattttac	186900
tcttccctcca	tctcaaaaac	ttttcaaaaa	ttttctgggg	tcgctcgagt	ttagaggggtg	186960
gaccgctggg	tcggtctgat	gtcaagtttg	gaggggtggac	cgctgggtcg	agccaatgtc	187020
agattgcacc	agaaatgaca	gagctccaga	aacgtacatt	aactttatac	tgtttctgga	187080
acatgtattt	ctggagactg	acacgtccag	aaagggtagg	gcatagtcta	ctgttttcac	187140
tagggaattt	tactctgcct	ccaccttaaa	aacttttcaa	aaatttttct	gggtcgctcg	187200
agtttagagg	gtggaccgct	gggtcgggtc	aatgtcaagt	tcggaggggtg	gacctctggg	187260
tcgagccaat	gtcagattgc	accagaaatg	acagagctcc	agaaacgtac	attaactcta	187320
tactgtttct	ggaacatgta	tttctgggca	cggagcacc	cagaaaagggtg	aggctgttaa	187380
cacccttaat	ccactagaga	attttactct	tcctccatct	caaaaacttt	tcaaaaacttt	187440
ttctgggtca	ctcgagttta	gaggggtggac	cgctgggtcg	gccgaatgtc	aagttcggag	187500
gggtggaccgc	tgggtcgagc	caatgtcaga	ttacaccaga	aacaccagag	ctccagaaac	187560
gtacattaac	ttttacgtgt	ttctggaaca	tgtgtttctg	gagactgacg	caccagaaaa	187620
gggaagggta	taacctacta	tttacactag	ggaattttac	tcttccctcca	tctcaaaaac	187680
ttttcaaaaa	ttttctggg	tcgctcgagt	ttagaggggtg	gaccgctggg	tcggtctgat	187740
gtcaagttcg	gaggggtggac	ctctgggtcg	acccgaggtc	agattacacc	agaaaatagca	187800
gagctacaga	aacgtacatt	aacttttacg	tgtttctgga	atagacgttt	ctggtcatgg	187860
acgcacccag	aaaggggaagg	gtgtaaccta	ctgttttgcg	tagagaattt	tactcttctc	187920
ccatctcaaa	aacttttcaa	aaatttttct	gggtcgctcg	agtttagggg	gtggaccgct	187980
cactcggctc	aatgtcaagt	tcggaggggtg	gaccgctggg	tcgagccaat	gtcagattac	188040
accagaaaacg	tttgttgctc	cagaaaacgat	aacacacatt	tctggagcag	tagttagtta	188100
tttctggagc	acgttaaaat	aagtagggag	aataatacat	aacattatat	tttcaattcc	188160
attttattgg	ttctatccat	ctaaaggtgt	tagaggaaat	actttcattg	tagtataaga	188220
atccttactg	aaatttttca	aatcatccaa	agagaatata	atctttccag	tataacttag	188280
agggttctct	gggtagctgt	ttttcgtgtc	ctggcgaga	cggtagcttat	tattgcttcc	188340
catctttaaa	tcttgccttc	tactcctctc	cagatgcccg	aataaggcta	tggatgaaat	188400
aggggtcttt	cctacaggtt	cgacaataag	tcttctata	gatgcgttcc	gtatcaagtc	188460
ttgcttgcc	ttcttcacat	tcctgtcgga	tgcggcttta	ctcccactag	ttttccatat	188520
ctctgattcc	gtttgcatct	ctctaaatgc	tttcagatc	aatttttctc	tgtaaacatt	188580



ttccactttg	acataattat	tcttcaacat	gttatccact	aacatcttgg	atgaaaatgc	188640
agataagaga	ggtaagggcc	tccctggtac	aaaactagtg	tttgacgcat	caatcacatt	188700
tcttcgtgct	ctaacagcat	tccttaaagc	ccaaaagcgt	gcaaaggata	ataatgatga	188760
tgtattttca	tctgttgga	catttcccca	taagttatgt	gtaggaaact	tcaaatctcc	188820
ccatcttgcc	aatagtaaat	atactgctct	gtcgaaattg	ttcgttattg	aaggattgga	188880
gggtggctgca	gccgcgtaca	acaggtagaa	agaaagggca	gcaaaccctt	tagatacatg	188940
gacataatatt	gcagcttcaa	acacataccc	tttcttatcc	actaggtgca	tgttacgcac	189000
ttcttcgtcc	tctctttttt	gtttcatttt	tccttcaaaa	cctatatcac	cacctaaccc	189060
ttccctctct	aatagaccgg	aatctgcaga	gagtatctca	accatcaact	cccctaacc	189120
attgacattt	tctgctcctt	tattattacc	accaccacca	cctaaagcat	cgagtactgc	189180
tttccagtag	gtctctagtt	cctgttctcc	catagacggg	ttactttgtt	cctttaaaact	189240
gatggtaaaa	gtgtctagag	attttgttat	gaaaccagat	tcacaacgct	ggcagacgaa	189300
taaagaatgc	atatctttca	aaaacacttg	aagaattaac	ctgcttttca	ctgcagagct	189360
attagttttc	tgtgttagcc	ctccagagca	taaaaataat	gcattccata	gcacatacat	189420
tgaaaatcgg	gcagctgagt	cgggatcaag	aacgcgggta	acagagtagt	tggagaacat	189480
aacttgctca	ttattgacaa	gaaaatactc	caaagtttct	gggtgaaacaa	tatcgttaaa	189540
cttgttacat	ctcttgattt	cacgttgaag	agcttgaagg	gcggtaggtg	tgctgttagg	189600
ccctaaaatt	ccgtctatcg	cctttggaag	atatgggtgg	tccttaaaga	gattctcagc	189660
agtcataatcc	gtagttttta	ttcgttctct	ttccctctcg	accattctct	ttccctttc	189720
tgtgacactt	tcggatggca	tctgcagatg	gggattgttc	acataattat	ccagtttagt	189780
tgtcaaagga	tttggtttag	cgtttctggt	cctggacagc	actgccctgt	actctcgctc	189840
gatggctgct	gataatagta	ccatgatttc	cctccctaaa	agagcttcat	ggaaagaacc	189900
taccttgggg	gggtagaatc	cttcctgtat	tacagaaacc	atgttggtta	aaacagtcga	189960
tatagttgga	acaacataag	attgtccatc	ttcttccactg	actgtaactc	catcactaaa	190020
ggacccttca	ataaagtctt	ttatcctttt	ttcaatcatg	tcagcagcgc	gtccgtactc	190080
taggtgagga	ggcatggcag	caaagggatc	taaacccttc	tctttcattc	tagagcgcat	190140
agattgcttc	catcttagaa	tatcacaaag	agtatatcct	agaaatagag	caatggctct	190200
ttttgctcgt	cccattgcca	gatcggccac	cgttgaattg	tttctgttgg	ccaacatttc	190260
tgaatcaatt	acccatgaag	gttcattagt	atcacctcta	gtggaataaa	gaactagatt	190320
agctgctgat	ctaaacacac	cgtccataat	attctcaaaag	gcattcaaaa	ggttcaataa	190380
gggtgcttgt	ttcaaaaaag	taaagtcaca	tagttcttcc	gctgcagttt	cgttcaatct	190440
ccaagagggg	tcaattatat	gctctctaac	acctctctata	tctctgtaga	cactcctcat	190500
gacggcatct	ttagatagggt	gtcctttggc	taatcttagt	cccgaattaa	aaaattctgt	190560
cgttttattg	aaccacgtgg	gtccgaatac	gtcagaagaa	ataacaacct	tcctttgttt	190620
ttcaggtgta	ttttgtaaat	ttgaaatggc	agcaactagc	ccttccgata	cataatttag	190680
tcgtttttca	tcggttttag	tttctgctac	taatgccaga	acatagtcct	ttgactctgt	190740
catgattttt	gtggtgctgg	ttaaagttagt	tttgtacacg	ctgcaatata	ttcgacctta	190800
aatgtcttcg	gttaaaggga	tatttatttg	ttggtactgt	tttcgtcaag	gctgaaaaat	190860
aattttcaaat	atacaatgga	gtacatggaa	gaaggagaca	tcgctgaaag	gcgctcagaa	190920
ggtgtcgact	atattctgga	cgaaaactct	gcttgtgtag	ttaatgtgaa	gagtatccgt	190980
aacaggctcg	gtgccatgga	cgcggaggag	gcacagtaag	cacaggacat	ttccgcccac	191040
cttgctcacc	atattatccg	tctggcccac	tgctccgaat	ccaacaagat	taaggatacc	191100
attgccagta	ttgcgggtct	tttcatcaac	aacatctttg	acaacaattc	aacaagaac	191160
aaacttaaaa	cgtataatca	attcaaggca	gagtcacaaa	acaagtctag	cgttctcaat	191220
atctttggct	ctctagatcc	tctgagtatg	ctttctagct	tcatgggttc	tgatccagca	191280
aagagtggag	gggaaaattt	ggacaaatct	ttgggtgtgc	tctttgaggt	gcttcaaaa	191340
tacaaccctt	gcaagattga	cgatattgtc	cttctggaaa	tgtgcccata	caagtgcgcc	191400
gcctgcaccg	gtctcaagga	agccatccgc	caggaacaac	ccatggaagc	aatgttgttg	191460
tttttcaaat	gtatcaacca	taataggttc	aattttggaa	gcgacataaa	gtcagcatac	191520
gcctctgaaa	catgcatgag	atactctcag	gacgaacgcg	cagtcgttgt	gcctctgagg	191580
agtatcctcc	tcggctgcct	cgacagggat	gatccagctc	atactctttc	ctctttcggg	191640
gatacatctg	agtatgctga	ttcagataac	gcttgggttt	cttccctggt	tgacgccgtc	191700
tactagaatgc	ctatggtaga	cagagccgtt	attgtcact	tttacgtgta	cacaatgctc	191760
agccgacata	ggcaggtatc	tggagacagc	ttcaagcagt	ttgtctatac	cgtatttgtt	191820
cgtatgattt	actctgcgat	tgaaattttg	ttctgtgata	ctgaaaattc	gtctgtagaa	191880
tgtgatggaa	agcacttttt	gagctatggt	aatgccatgg	ttaacgtatc	cgtgctgggt	191940
tctactgtta	acgtactaaa	agcctaccgt	tcagtgggtg	tggatcaagc	atccgtcgca	192000
cgggttctag	acattatttc	cggaggatgg	aagaagaact	acccctcacc	tgaccacatc	192060
aagaggggtg	cgtacgacat	ctctcaagtc	atcaatcacc	ttgcatcacc	ttctagaatg	192120
gttaaaggta	acaacaaggc	tagcaacggt	acatctggcc	tggatagtat	caggtctggt	192180
cgtcaagcag	aaaaatatat	cccgtttgga	atacttgaaa	ataaagcagg	gtatggtgta	192240

ataaacattg	ccaagcacia	catcagccgt	ccagcaagag	aacaatccaa	cggaagggaac	192300
tttaattgca	acgctttaca	cattctacct	tcaattaagg	gctgtgaggc	acttgggggca	192360
caaaagggga	gcgcagatca	aactgtcaat	gttttttgata	atthttgtcgc	atctcatatg	192420
gatattgcca	tgaaaaagca	ggggtcgggg	aagattcttg	gactgtcac	tagcatgatt	192480
gacaggcaag	gtctgactac	ttcattccct	agtagtgaag	cggaatacaa	gaagagaatc	192540
catgatttca	caagatacgt	catctttctct	tcaacaccca	tcaacgacga	actagtcaat	192600
tctcgctgta	ttcttcccca	ttctaattgtt	ctgaactccc	ctatcagctt	gagaaatatt	192660
gacccagaat	cagtccccga	cactcgattc	cactttctgc	tcatgatgtg	gcagcggcca	192720
aatatcgatg	aacctaatct	ttctgctcta	actacaagtc	agctagaatt	gttgcttagc	192780
aagaaccaa	aatgggacaa	actcaccacc	agagcgttct	tcaatatcga	caggatcaat	192840
ttccagatgg	cagagcgtat	cattaagaac	gtttctggaa	gcggcttcct	agatgggagt	192900
aaaactgcct	cttcttcttc	ctcagcgctt	aactttttcc	aaatcttcag	tggtgctgaa	192960
tgactgcaa	agcagctcca	aagtattcgc	aaattcattg	gagaatctat	gcagcatgta	193020
caaaaggaat	ggagtagtgc	agtaaacaat	gggaacagag	gagtagaaaa	ttatgacgga	193080
ctcaatgctc	agttctctga	agaactgttc	gagctgctct	acaaattgat	catcgaggag	193140
gatatgctgc	catccagcct	gatcgctca	tctgaattct	tgagcaacta	cgtcaacgcc	193200
atggatgaac	ttcttatcag	agctaattgct	tcttagatag	ggtttttatg	aaataaaaca	193260
aatataaatt	atatatcatg	caatttttatt	tatacctttt	tttgtatgtg	atacaatatt	193320
tttagtgta	tgacagtga	accattatat	tattatctgt	aataggtcct	ggtaaatttt	193380
taatgaaag	aagattaggt	gcacaatcta	ttgtcataaa	aatgtccagg	tcaatttctt	193440
caacaccttc	ctcaaagtta	tagaagaaac	ttgtttcatt	atthttatta	gcaataaaca	193500
agtacgtttt	ttccttagct	gtgcttgca	agcttatcca	atagtctgat	gcctttttta	193560
cattataaca	gttggggagt	aaactacaaa	aacagtcctc	tgthttgttt	gtagcgtata	193620
cccgtacacg	gttaccttct	actacaaatg	aaaactctat	ccgggctttt	tcaatacgtc	193680
ctcccacagc	tgctctacct	tgthtttgta	gcacgggttc	taatccaact	tgtagagaat	193740
ccagattcgt	tgctgcacta	ataacattaa	atgtcctctg	tttcatgttg	acatgtacag	193800
acattcgtcc	ttgccaagtg	cgcatacac	gatcactgat	aaaacctcct	cgtatattat	193860
ttatcaccaa	ttcctttgag	taccgggcaa	ggtaatcagg	tgtaggttga	ggaggttggt	193920
gttgtctagc	tagttgtgaa	gaagggtgtg	ttgtcaacat	caccatttct	gcgtctgggt	193980
ctgttgatct	ctgtaggctt	tcacccccat	tttcatcaga	aggatcaact	gttactgtac	194040
caaacaataa	aattttgggc	aagtctatta	cgagttgttc	gtaagtata	ttggtagatt	194100
tttgtgcaac	aaattcatcc	atcgthtctt	tgactctgct	atggttcgct	cctatgagat	194160
cgaacaagtt	ttcatcagca	attgcctgtc	cgacatcggg	gacgttttat	tcaccttcag	194220
ggattgaaac	ttttccatct	gttattttat	cctttattaa	ccatacgttt	tctgatttgt	194280
atgtgtcaaa	ttcctgatct	gctaattttg	tgaagaattc	atcctcagct	ttttcaagag	194340
ttcttgcaaa	atgctctctc	gtttgacgtt	gaagcgctgt	ctcattgata	acatcaactg	194400
ttcctcccc	ttcctctcca	ttattattat	tatcgctgct	agthttttca	tcacgtcag	194460
tgtattttaga	gtctaattta	tattttcttac	tcttttttgt	tttggcattt	gaatgtagag	194520
ttcccttccc	tacgcccag	atagcaaaaga	agagggcaat	agaagctcct	attaccaaaa	194580
caataaccgc	aaaaacaatg	tatattcctc	cattgtttga	aagtgccatt	tttaaacctt	194640
tcaattctta	tctaaaaaac	aaacagattg	aaactctatg	ataaatcacc	actcgcgttt	194700
tcgatttccag	gcattgtgaat	tgtaattgaga	gagccggaat	cctcgatgg	gttgattatg	194760
ccatttgtat	catgactaat	atatatagga	tcatttctct	ttagataatc	acctatgtta	194820
tcagaagaat	cactatttct	tagatttgat	aatatcacta	cactttttgg	tctgtttaca	194880
ataccagcga	acgtggattt	gcggttaaat	tctaactgta	gagttttggc	tgaagaatct	194940
attttgaacg	ataaacttaa	atcaatttaa	tccccttgag	cagctccagg	gggagtgaat	195000
gataacatgt	gtttatcggt	tattggtaca	ctaaatgaag	ttgaagagtc	ggtaaaggca	195060
atagtattgt	acttgaaaga	aaatgatgta	tgataaaaa	caggatatata	tcttctccac	195120
cgctcctctt	ttaaaatttc	agttatgaaa	tcttccctgg	gttccatgat	ccgtctattc	195180
acatcgtata	ctttttccat	aaccactgg	attcctttct	caaaaattgc	ttcacagata	195240
acaggattga	aagcgtcttc	gtaaaagatgt	tcaaaagtca	ttacattagt	taaatggcta	195300
ttgatttctt	gccttacttg	atcctctata	tttttatag	aggcattttt	tgctaccgaa	195360
taggcaattg	cacctaaagc	atactttaat	tggtcactct	ttcttgatg	ccaagaattg	195420
tcatggtgag	cagctcccaa	ttgagcctcc	actgaagcga	tttcagcttc	agtcacagca	195480
ggctcatctc	cgtagacga	cattattttac	tcctccttaa	aagcagtgat	tgcaaacat	195540
atatcgtccc	ttgtatcgcc	tctttttata	cgthtaacag	cgtatgttgc	agattctccg	195600
tttagtatat	gggtgtgttt	atatctagat	ttacatacaa	tatcaatata	ttactcttc	195660
ccttcaacaa	gaatattgact	acgcttctc	ctgttctgatt	aaatgcttca	195720	
aagcttagca	cgctattaac	tacgtcacta	cgagacacaa	gaatcactgc	atctacatcc	195780
tcatcattta	tggaaacgaaa	caattttgcc	gttattttaa	aatctcccaa	gggaatggag	195840
aaatcgtacc	ttgtaaattc	tccatcgta	ccaccaattc	tcaaatgtga	cgtagcaaca	195900

ctacaccgcc	ctacttgacc	agtttcttgt	tgaatTTTT	cgtctatata	tgtgcaagat	195960
attggggggc	gataccaatg	tacaagattt	tgcgttagtt	ttcgttttagc	caagtaatca	196020
gttattattt	gtctaacatc	tattccttgt	ttttcaattg	ccttttaaaac	acaatccttt	196080
tttatgtctt	ctaggcgaga	tctcacatcc	agttctgcat	ccttcattgta	aactctgtct	196140
ccatgtctta	cagtgttctt	caacaattcg	tttatgtgtt	tggtttatggg	gtgtatcaac	196200
aattcagata	gattagaaac	accacgtctt	tctacaattt	tatcaaggat	gaggatcaac	196260
ctctcatcgc	ttatagtgtg	gccattagac	atgctgttgc	tacttgcttt	tggatatacaa	196320
catcaacagg	tgtaaaaaac	atacaaaaaga	aataataaaa	atacttttagg	ggaagaaaat	196380
gtttagaagt	attttcaacc	ttttatttgt	ctggggcgaa	gattatatac	attcaatctt	196440
ttaacgtaga	ccitttctcg	tcttcttctt	actcctactt	cTTTTtctcc	tcctttctgg	196500
tcctggaact	ggaactgggg	ctcggttaact	ggagcaagtgt	tgttggtggt	agtgcagat	196560
attacctcct	ctatttttgg	tatcctgagt	acaatttttg	gttttggttt	taattgtact	196620
gtagagggag	ttttctgtgt	tacccttata	gcctgtttca	ggaacaagggt	taggtgtcta	196680
taccaagggt	tagccattgc	tcaagtgtatg	ctccaccccg	tgtttttccc	ccttatatac	196740
aattttctac	gtggtcaata	aaggaaacac	aaaaaatatt	catcaattct	tttatttgat	196800
caaagaacat	gttttacaca	gtttgggcac	agataatcag	ggttaagtct	caaaccocat	196860
tgggttcatta	caatgtactt	gacaatatca	gaatggtcag	atTTTTgaga	agcattagga	196920
agagctaaat	tggtagagtgt	tctaagagca	tgaacggtt	tattcacaac	cattgccaga	196980
aacatgggaa	cagatagtga	agaaacttga	ttctcatccg	caatatcctc	tgtatttgcc	197040
ctcttgtggg	atgaagagga	caaaaacaca	gggtatttca	gattactaac	catcaattgg	197100
tacaccatac	aggctaacat	tacagcaaca	taatcaacac	cagttttatc	actctttagt	197160
gtttttaatg	aggaaagtct	cctcttacct	tcagtatact	ggagagccat	ttcaaacac	197220
atgtgtaaat	ggtggcaaga	atacatcaaa	tccttaacat	ccattgaacc	ccaggcatca	197280
cacaaatggt	ttgttagtga	agaaccaact	cctattaatt	ccccttcctt	atcatatcct	197340
acaacatcca	tatcaggatt	tggtaaactc	ataccaatac	aattgtcaaa	attgatggct	197400
ggcctattct	ttgtggcagg	agagtgcac	ttgatgttaa	aatctctccc	aaagaatctt	197460
gtggtcttaa	caaagggcat	ttctacaaat	aaaactgaat	tgagcatgct	atttagttga	197520
gcccattgcg	cttgtctaga	acgagcaata	ttcatagcca	ttgtattcaa	agatttagga	197580
ttaatctggc	acatttgaag	aggtgatggt	aatttattgc	acatggactc	gtaatttgct	197640
ggtttcttac	ggcgcaattg	ttgttccctc	gcatcacaa	tattgtcttc	ttctacttca	197700
gtgatagtga	cagaagaaga	agaagattct	ccgtaatttg	ttgcagcctt	agcacaaaga	197760
ctcattgaag	tgtacaatat	gaaaaagggt	ttagtattct	cctctaaacg	ggaactgat	197820
tttacgtctc	tcaattctag	gaccttggtt	ctgatcaagt	ttttcattaa	aactgatgga	197880
agaacggatg	atacgtcact	aattctatcc	ttctttttgt	tgtcactgga	atgccacta	197940
aagtcatctc	atgctttctt	gtgatgatta	aattcgtgtt	gacgctgttt	tattccatgc	198000
tttccaaatt	tgatacatct	ttccatcatt	aaagatgaat	ggatataact	cattcctcta	198060
ttcaaaagcc	cgataatttt	cttaacatcc	tcattcttat	tctcaaaagac	aaatgttgta	198120
tcacaaggaa	aaacaataag	tcttttagac	atagcttcgt	cgacaaagaa	tgagaagatg	198180
taactgttgg	acatcataat	aggagacaca	gtaatgctgc	tttttgccgc	ctgttgga	198240
atgcctctaa	tcattgacgt	gttttggggt	gaagtcatat	tctttaataa	actggaggga	198300
aattggcccg	gaagcatatg	tttgggtata	gaactggaca	catcattctt	ttgggggatg	198360
gccttagcga	tttcggtaat	aatggaaaga	gatccgtcac	tgttatgttt	ttggagtgca	198420
gacaccaatt	gccttctttt	aggcgtatcg	acattgaaac	aagcctcttt	attactgtct	198480
cctccaccac	tagagttaga	agcaaaggaa	aattgtgaac	tttgttgcc	attttctcca	198540
aaaacactta	aaaagtttgc	aatatcactg	ctgttatcgt	ctgttttagg	taaaagaag	198600
gtattttttg	atgaattaga	ggaagagtcg	tccttttctg	cagaaaagca	atttttatcc	198660
atacacatgc	gaataaaagg	tcagaatttc	ttagccgaca	tcttatcctt	ctcaatcgc	198720
ttctttaagc	aagaaaggac	attttcgcag	tcagaagtgg	aagttgacag	aaccactgaa	198780
ccatcagaag	taaatccagg	aaaactatat	tttgacggtt	tcttaaccac	atcgaataaa	198840
gatgcgtcat	aacgttcgta	catgttcata	aagtccagat	ctctatccac	taatagtttt	198900
gcagcttctg	ggctccttat	tatggacgat	aacttgcgag	aaatagtctt	gatgttacga	198960
atactgtcaa	tattaatatt	cgacagagta	gggcccagaa	gagccttctc	ggcagcatca	199020
tcgacagcta	caaagtgttg	tttcttggga	gggggaagta	aaactgggta	gtttgttttt	199080
gaaagatcat	cagaaatttt	tccagtgtaa	acgcaatggt	acgcaatcaa	tacaccagggt	199140
gcggatgtat	taccatcaat	aactgacttt	ttaacgcccc	cgttagtcga	caggcaatta	199200
tccccagggt	atttttta	ttcgagcatg	tcctgggcag	acaaagatga	cgtgatcatc	199260
tcggaaaaac	tttcagtggc	agggttgcag	actaaaacag	gcgaaccgtc	gcgatcaaat	199320
gatgtattcc	atttcaaaag	gccattgttt	ttctgtatta	tgttcataat	cgcatacaat	199380
ttgtcttcaa	ttttcatatt	ctgcttattt	ttcttgttct	tgtctataaa	tttctcatte	199440
cattcatcta	tatcttcttt	acaatcgttc	caggaatgaa	gattattaac	ctggcagttt	199500
tcatcaataa	tggattcaaa	tgattcgtta	tcgcatattt	tggatccaaa	tttaccta	199560



cggccaaacta	ttctaccctc	atggcctgta	tttgccaaaa	cactgccctt	tccttgcttc	203280
ttttccgaag	aattgacaac	tttcttgtag	aagttaaacc	aaaacgagca	atcacttcct	203340
tcatgtttca	attggtcttc	agaaatttca	ttggtctggt	tagcttcatt	attattatta	203400
gagagaaggt	gacctgatga	ggatgatcca	ccccacacc	caccatttcc	tccttcgcca	203460
acagaagaca	tgtaatcgcc	aggacataty	atgttcatat	cataaacgct	tccagtatca	203520
gtcaaatcgc	ggcactcttt	tcttatcttg	ttagtctctt	cctcaactgc	cttattccac	203580
ctttcaatgg	catccattcc	agcacttaac	gcagtttcag	tgacaccgta	caaagcatac	203640
tcttttttcg	agaaaggaac	acaaatctta	aaaaagtctt	caatgtcaga	aattactttc	203700
acaatgtcat	tatttatctg	attttgtgtg	atcatgttct	taattcgggt	aagtacatag	203760
ccatcgtaac	aaaacaagcc	actttctctg	tcttctctgt	ttctggcgtc	catgtgtatc	203820
tggctgaaac	ccgtgcggga	tctcgagtcc	tcatcaacta	cggatgatga	tccccgacta	203880
gaatcatcat	tgtaaaaacc	acccgcgtta	ttaaaaacct	ctccgccaga	tttctttgat	203940
cgtcccttat	tacttccacc	ctttcctgtc	atttgttttg	caccactatt	atcctccttg	204000
ttaatctctt	ctagagaagg	atcgataaat	gcttcttcgt	tagcgtccat	tactttagct	204060
tgaccggcag	gaatgltgtg	gaaagaaatg	ttagtgttga	tgattctaga	ggacttttct	204120
acagcagcga	gttgggcagt	tgtacaggaa	ttactaacca	gattgtagtt	agtatcgatt	204180
ctagtctctt	gcagttcttg	gcgagtcttg	gcgcggaact	cgtccaaaga	tttaacactg	204240
acatcattaa	cgtccaaaat	acaacaatcc	aacaaatcat	ttgaagaagc	acgcacagag	204300
tctgccgctt	ttactaaaac	aagccttgca	atacgtttca	ttaccttccc	aacacgtttt	204360
tcatacatat	cattgctttt	ttcgctattt	ttgtcggcc	aagaggaaac	actaaccaaa	204420
ttcttaacta	gtgaaagggt	cctagaaaga	agcctgcttc	gccgtcccat	taggacgtga	204480
gaaaatctat	aggtgataaa	ttgcaaccaa	ttactagtat	caaattccatt	aaaatcaaac	204540
gacatgagaa	tctttcgggt	catttcttca	atgacaatat	cattacttgc	aacttgtcca	204600
ataacataty	cggcaaaacg	gtaaaagtgg	agaggatctt	cttcaatatc	gtaaaatcta	204660
ccttctttta	tgctagcgtc	ggcctttttg	ggagaatctc	ccttgccctg	acagctgcac	204720
cacttgctag	tgatttttaa	gcacgcaatg	agtggtacca	ttttatccac	agatagacgc	204780
aaaggagcag	ttgcaatttc	tcgcacactt	tctccctgca	tcatgatacg	atacatagcg	204840
accatttgcta	gcacacaaga	gagggacatt	tctgtcttga	gaacctcttc	tgttgcagcc	204900
atccactctg	cccgaactcg	caaagtagac	ttgtttcgtg	gataaatattc	gtcttcaacc	204960
caaatttgat	gcctaacatc	attgctaggg	gtgggtttta	tottcaattt	cttaattacg	205020
ttgttatatt	ttagcaagtt	gcttgaattt	atgaaatttt	caacagatct	aatgtcgttg	205080
tctttgagag	caaactcata	ggccttggtg	acctgcgttc	tccagtcctt	attgtcgtca	205140
ttcaaaagac	attcccacgt	atcagcatcc	aaatttgtgt	ccagatcttt	accaaaaaat	205200
gccctcgcg	taacgcccag	aaattctgat	agctcaata	cggcgccaga	ttcgtccaca	205260
tttgacacac	ctgcctttta	atttagcatg	tactaccgt	cccaatcctt	aactaaactt	205320
ccccaatcaa	taggaaggct	gttattaaag	cactccctaa	cagtccttag	aattgtccac	205380
agtctagaag	ctggactatt	ggttocagaa	ccgtatcttg	ttatcaaatt	tgaaaagAAC	205440
tcagaatcat	tacttgtttt	tccattcaaa	gctgtacgca	tgatacatte	ttccatcatt	205500
ttaaagcact	cgtatatgtt	aagttttgaa	acgttatggt	ccttgaaata	atttgctgca	205560
tctaaacaca	tctgttgacg	tttttcttta	agttccttct	cagaaagcga	ttgattaaac	205620
acttttagttg	gttttagagg	atcaacaact	gtagaacca	gcacaacaga	ctctaaatct	205680
tgatttcgac	acaatgccca	aataccagaa	aagaaaccat	tattaacact	cttatctgat	205740
acaccatcac	ctcccgatga	acatgccaat	ctggctaaac	ttctaggtct	aaacgtgtta	205800
atgagtttgt	tatcttcact	gatcacattt	tcatagatac	ccaaaggtaa	tgctgaagat	205860
agacattttc	ctgccaattt	acactgtttt	atggcatccc	acctcttatt	atctaaaata	205920
gacctaatat	taccattaac	tttgtccttt	ttcacaaatc	ctaaagctgc	agtaacatca	205980
ctaaagtatc	tcctttcaca	tgccatttgt	ttgaaacgga	gagctacagt	cagatctttc	206040
agatgttttag	aactaggaat	tgggagtagt	tttatgcact	tttcgttaga	catggtattg	206100
ttactactaa	ggggagaatc	gctcttgaag	ttattaagag	aagcacaaga	acaaaaggga	206160
gggatattgt	tgttgctgtt	agcgggttccg	tcttcaactat	tgacattagg	cttgacaatt	206220
tctcgtacag	gcatagtaaa	aggcaactacc	caattcctca	ttctcaaaga	tcttctatta	206280
atttctttca	aaatatcacc	agattcagct	gcagaattgt	ccaatgttag	tacatcacac	206340
aaatttttct	ccatttcctc	cacaatgaca	acttttggtg	acttttggtg	atgacataaa	206400
ctggatctaa	agttgaatgt	tttgatgaaa	ttaacatccc	taaacacatc	atcttgtagc	206460
ttattaatac	tactagcatt	gtcagatggt	acattttcat	catcactaga	accaaattcg	206520
cctagaaaac	cataacaatc	atcttcatca	tcaccaccat	aataatcacc	tgttgcacct	206580
cccttggcag	ctccataatt	ctcctcatcg	tcactctctt	catcactttc	ttcatcgctt	206640
tcgttactgg	ccacattgtt	caggatctgg	cgagcagaaa	cgtcgtcgtt	gtcacatttt	206700
cttagatatt	cttggtttttg	ttgatagagt	ttctccctct	tctccgactt	cttgctctgcc	206760
aaaacagctt	ccttatcaga	cttgacgctg	gcaaaaagtc	gttggaaggt	ctccatgggg	206820
gcgcattcat	acattggtcc	aaaagacttg	tcagattcat	cataatcaaa	cgttttaaat	206880

gtatcgttgt	tcaattcgta	gaccattgtg	ctttcatcat	acttaatttt	gcacaatttg	206940
ctgtttgtct	tgttcttgat	cctctccatg	gaattttcgt	atgctgcaac	tcccattctg	207000
cagacagatt	gcttttcaag	cacacttttg	ggtagacata	tgctaagttc	tcgggttaatt	207060
gcgtcgca	tgtgtctaaa	cttctcaaaa	cctccttcaa	attcaaaatc	ttcacgaatt	207120
atttgcgtgat	agtactcttc	tggatgagaa	gtgatgttgt	acagcgca	aaaccgctca	207180
gcaagagtaa	cgactgtcca	cgcgaatatg	gtgtcgttct	tgtagcacc	cgtggcgaa	207240
gttgacgtga	tgatggacaa	gagcttcttg	atattttcca	aagaagctat	ctggttttca	207300
tcagatgacg	aagtggatgt	tgcaccttct	tctagactag	aaaccatctc	tggtgttgtt	207360
gatgatgttg	atgcagcagc	agcagcagca	gcttcttctt	cactttgagt	aggtggagga	207420
ggaatactct	ccaattcttc	aattgttaga	ggtttttagt	attctttgtg	ggtaaatg	207480
cacatttctg	caaatttgcc	ctttgatagt	gcaccccatg	tggtattctt	gagtccttgt	207540
tgtagaatgt	ccgtatctct	gggcaatttc	aggtttctgt	tctcatatct	cacaacaatt	207600
tgaaggatgg	cgaaaggcat	gattgtgtca	aacttgatgg	tggttaacgc	agtcaaatc	207660
ctgataatgt	tctttgtttc	tggtttcacg	cgttcaacac	aattcttctt	gctcatggct	207720
ttctggggaa	gaagaaactc	cattgccaga	agaggagagt	tgtcaatgtc	actcagaaag	207780
aaaccgtgtt	ttctcttgaa	gaaacacaca	atctgggtcaa	gcttatcatc	gtcatttgatg	207840
atgctattat	agacattggc	agcaagggga	gggatttgtt	taggggctgg	aggaggcttg	207900
ttgatgttgt	tctcgtctcc	agaaccgccc	ccagatttcc	ctccagatcc	tccttctcca	207960
ttatcttctt	cctcgtcgtc	tccatctcca	gaatcgtcca	atcctttaca	cgccatcttc	208020
tcctttaata	gggtagcatt	tgtcttggtc	aataacaagc	cacctcttcc	tgccagaaac	208080
ttgtaacaca	tgccaattgt	acagaaatct	tgatccataa	tcaaactttt	gttctctgtaa	208140
ggcatttctg	gccaagtctc	agaaaaaagg	atctttcttc	tccaatcgtc	ggtattggcc	208200
atagaaccca	tgagggtttc	gagatataga	cactttttct	cgaaactttc	cagtgttttg	208260
cacatccact	cttcggctct	atctgccttc	acgacaccca	gagaagtcac	catgttttcg	208320
ttgctgatgg	agagaatggt	cttcttttct	cccaccaact	gcgcattggt	ataggggaa	208380
gcagtgtcgc	tcactttaca	atgtccggga	agattccagt	tatcactccc	gccaatatca	208440
atcttgaaac	aggcaacagt	tttcttgatc	atactttcaa	cacgttcgac	gattctgggc	208500
atacattcac	tcagtaattc	aagttgtacc	tttaagattg	actttaacaa	caccacattt	208560
tctgtgtctg	aaacagctgc	aagacgttcc	tcttcaagga	tggttgggta	gaatattacc	208620
ttgagcccca	tcaaagaggc	attggtgatg	aaggcaacct	taaacactgg	cacgtctgtt	208680
gctgttcttc	cacacccag	aaactctctg	aaggcaggag	tcacacatag	tttctctgca	208740
ctagaagttt	cttccctata	agcctctatc	ataataataa	gttccctgag	aagctttcca	208800
gttccaacgt	ggggctcctc	ttcatccatg	cgcttatttt	cattctggca	agaacgctca	208860
ttcagcttgt	ccagaaattg	atcaccagta	atttctggaa	ccttttttac	tttgtacatg	208920
cctttcttct	tggtttttgc	aacatccttt	aaagtgtcag	gttgactgta	cattatcaat	208980
tttccatcgt	ttcctctgtt	cactgtgtca	tattcaacaa	cattacatcc	agcatacaga	209040
agattgtttg	ggttgacact	gcggtctccg	ccgactctat	cttgtagcgc	tagatgcccc	209100
atttgtttga	ccagagggtt	gccattcttg	gtcaagttat	cagtgactag	aatttgttcc	209160
ccttcggtga	gaaggagaag	aggtagacgt	ctgtgtcgt	actcgggcag	tttgttgag	209220
agacagatgc	gtagagtcgt	tggttctctc	ttcatctggt	tggcacattt	tattaaagtga	209280
tctcgcacat	cactttcaag	aaacttgcct	cctcctaggt	tgagcacatt	tgcccttaaac	209340
atattgacga	gtgcagtggt	ctggttataa	gttttgtttt	ttcttggtta	gttatttttc	209400
tgaatgtagg	ataaaaataa	gtaacttttc	tttgtgtcgt	gttcaccaag	gaatcttttg	209460
atatcaatta	cagctacaac	tggttaaggg	gttgtccgac	gaccgtctgt	atcctacggt	209520
atccagtgtc	ggactgtgag	gcagagagct	tgtgagggcg	tatttataca	agaatagcag	209580
ctggtgccaa	ctagtgtatt	tatagaacag	aatcatgga	ccagtaccca	gaagtgaggg	209640
atactcctca	gacagaacag	gaacaggcgg	cggcgcaaca	acaacaagca	gcaacaacaa	209700
cagctgccgc	tgctgctgcc	gctgtctcta	cgcagtacag	taacactgtt	tctgcagaaa	209760
ctttatccgc	catttctgaa	gatggaaaat	tgtagagggtc	aatcgcagct	tcgtgctgga	209820
tcaataacct	taaccctgat	gaaaaaatgg	ctcaacgtgt	ccaattccat	ccacttagtt	209880
ccacgaccac	atacgattca	gaaaaatgtga	accctggtag	ttctgttgtg	tttttgaagc	209940
ctagagccct	cccacccggg	ggcacgtgtt	tagcccccaa	ctacattgca	gtgcctactc	210000
ttcgtgctgc	gtcagaaatt	atcgattcta	gtgtcatcaac	tagtctatac	caatgttcaa	210060
tgttcaattc	gtggaatctt	attcccattt	ttatgagtaa	tagcaaacat	agtcaattcg	210120
gtgatcgtgt	catcaagaga	tcaatgatcc	gtaattgttt	ttccaaacaa	aaaaatgtag	210180
aaaatctttt	aaaggaattg	cgcaggagaa	aggttaacgc	cgccaaggca	ttttctcatg	210240
cagttcaaca	gaaatcagca	gtaaatactg	cccttgcagc	atggaacgcg	ggaagtgtcg	210300
caaaccctga	aaaatttagta	gattttttgca	aattggaagta	cagtccegat	agaaaaataca	210360
aagcaggagg	tcttttttagt	gcgtcagcaa	cagctcagtc	acaatcaggt	acttcttctt	210420
cctcagttga	acatacttct	aatgatttct	tactcgatat	cctaaaacgt	cacaagggaa	210480
cgtcccttga	tttagattca	gcaacaaaca	catttgatac	tgctctttct	agggttttta	210540

ccgagtttaa	ggaacaggct	agagcagccg	tagatgcagc	cgcagattca	gatcacctat	210600
ccgcgtcgga	ccctattttc	tctattgttc	gacacaatag	caggcgtgag	ggtattctag	210660
attcagtgcc	aaatatcggt	atgttggttc	cacgatcaaa	atattctgtc	gcagagtact	210720
tgatggcaga	cagagacgag	tctgcagata	tcgctgctaa	gatagggaca	aagattgcaa	210780
cagattttga	agcattgaga	ggagataata	ataagagaag	agcagatact	tcagtcgagt	210840
atctcaaaaga	atcactagcg	gattctattg	aaaaaacttc	cattaaaaat	accggcgata	210900
ttaattcgggt	cacaaatatt	cccacagata	cagaggagta	tgaattttct	ctccacatca	210960
cacagctttt	cgcgaggca	tttttgaaa	ctatgggaag	tttattgagc	tgtgctttcg	211020
gtgtacagtt	ccccttctca	gatgaagggt	ttgctggctat	cgaaaggatt	atacgtaaaa	211080
cagatccctga	tactggaaaag	gtttcagaaa	tggacccttc	ctctctttca	gatcaatatc	211140
tcttgttgggt	aggaaatttc	caggtttctc	ccttccacgt	ttcagatcct	aaggatatcg	211200
ttttcggtag	acaagttacg	cctaatacgc	ctattttgag	tatcattact	agatctaaaa	211260
atgataagaa	tgaacgtct	actattatta	atttcaggga	taggcttctt	gttaacgata	211320
ccgtgttgag	agacgccacg	caaacgtct	caacttccac	cccttctcaa	agaagagtcc	211380
ccaccgccgc	aggtgaacca	aaaaagccaa	tgctctctgg	ttgcctacct	atcatcaggg	211440
gacccagggt	agtgaccagg	gaaagtgatg	atatgattag	tggtcttggtg	gggtgattgtg	211500
atatctcttt	gggggtctat	tatgccatgg	gttcatcggc	tgcagctatt	gccgctggcc	211560
accagagagc	tctcgcttca	gccgagtcaa	ttaattcacc	catgatgaaa	aaattctcaa	211620
agaaggaggg	aaagtataca	gaggaggaaa	agcgaatcaa	gaaggctatg	cgtaggaatg	211680
catatcggtt	tgctaggatt	ttggctttgt	tggggcaaac	tgatgcccg	tatgggtatg	211740
tcgaacataa	ttctacccta	gattcgttct	ggctctcgaa	tgccgctatt	cgtagcaagg	211800
caaaggagga	tgctcttagc	cgtgcagaaa	tcttggcagt	taggaaacaa	ttggacggaa	211860
aatgttcttc	ttcaagggat	gaatattcca	tggtagagag	atatcttaga	gactcatttt	211920
tcaggtcagt	taataggtca	ggaggaggat	atgaaatgtt	tgatcaagg	tttgatatgg	211980
gaagatttgc	cgactttttg	agtgacaact	ctgctgccag	gaacgcttgg	caacagtacg	212040
cagaagtaat	gagaggactt	tctaagcatg	agaaacgcgt	attcaacatt	gaaggctctt	212100
tcagtgtctt	aaattctttc	aagttccccc	ttgttccaga	acaggggcgt	aaaaagactg	212160
taggtggaag	gcataggctt	aacaatttga	aggcgcccaa	taagatcatt	aacggcatca	212220
cagagatgac	tctccagtca	gccatcgatg	gtactggaat	ctctgatatt	attgggttcag	212280
ttagcgatgg	gtggggaaat	actacagctc	agccgtctcg	cgtcaaagct	cttaaaacat	212340
tgccaattt	cagcggaaac	ggaaatgttg	tgtcaattcc	agttagtcca	gcagtaaagt	212400
gtgccgcggg	aagtcgaggt	ggggaaactc	tgcaagtgtg	ggacattcct	tcagtcatta	212460
ttgcaaacct	aatctctgat	aagagaattc	tagatcaact	ttgcggagga	gggatgaacc	212520
tcgctcacga	aatcacaaac	tttatcgaga	cgattgcagg	taaagaacat	acaggaaagg	212580
aatccgtttt	cttgtctcct	agattgtctg	tcattctttt	gcgttatatt	tggttcaacg	212640
cagcggttgt	ttcccttacg	gatagcaaca	ttaaaatgcc	cctcaacaca	atgtccgagg	212700
gcactggcga	tgacatttat	agggaactat	tggccatcag	aggcatgggt	aataattaca	212760
acagtgtctt	ctcctctatt	tcagtcaagg	ctatttccga	taggtacaat	tgcggtagtg	212820
gaaatacttc	aaccagtaat	aagaatgtga	ccattaagac	tcaaggtgaa	ttgttaactg	212880
tccttcaaca	gactgtctat	gccttgtccg	cctttacca	caagggaggc	gtgggtgcaa	212940
ccccgatgc	tgccaatatg	gccaacgtta	tttccccaat	tgctaattgcg	gatgtagtta	213000
agaacaccaa	cgtgggtggt	tcagggttag	ataggatcac	tgagaccatc	aacttctttt	213060
cattttttgtc	tcagatcaaa	acaatgaacg	agaacattga	agagtatctt	aggagatata	213120
ggctaggaga	aggactagat	aagaaagaat	tggataattt	tgtgtatcca	aatattgcag	213180
ctattgttaa	gcgagaattg	ggggtaaagt	gttccgcatt	gtccagtaat	ctcgatactg	213240
atcgtccaat	tactatcgac	ctgaacactg	aacagccttt	gatcgtaaag	gctagcaagg	213300
gttatgcctc	taaccgctac	gctaaattat	tcaacaaaac	aacaagaaca	gcagcagaac	213360
aagctcagat	ggagcagtat	aatgcacaaa	tggctgccaa	tactattcct	caattagtaa	213420
acaggttgac	catccctgga	tccatcacgg	cagacactgc	catcaatgtc	gttaaagctt	213480
tcacagaaaa	tggaagaattt	agtaacgcag	aaacacactt	gggggttatg	ggtaacgcga	213540
ttaatgaaat	gcaacctctt	ttcacggacg	gattcaacgt	tgcgaacaag	cggttaacag	213600
ttaacgtggg	ttcagttagt	aagctgattc	agaattgggt	aaccgtatct	ctcattcttg	213660
ctcactcaaa	ggctagcccc	tatgtcttta	agcctctcgt	gcaagatttc	gctaagcttt	213720
tactggcagt	cactgcagag	acttctctgg	ttgtctctag	gtcccagaag	agtttcttcc	213780
ccattcctcc	ttcagtattt	tcttcagggt	gtcttttcaa	aattgatagg	gaaatgttcg	213840
ataatatgaa	gacagattat	gtagtggaa	taattagaca	gctatctaag	aatgctaccg	213900
ccgccatcga	aaggtgcaat	gattccgatt	cagctgctag	gattgccaag	tcagggtgaa	213960
tttataacaa	ggatgttgca	tcaaccactg	cagctcccg	aacttcttcc	tccgctttaa	214020
ccttgttcgc	caataatctc	cagaaccttg	caaaggtag	gtctatggga	gctctcccc	214080
atttcgatat	ggccgtggta	ccaaaacttc	atggcatttc	tcacgatcaa	atgttccgcc	214140
tatctacata	ttatcagggt	attcataaga	tggaaactta	cagcgattgc	aaaccagaag	214200

aatgggataa ttctcttctt ggaaataggg ctagcaaat ctttggcctt tcttcgggtga 214260  
gcgataacaa ccgttcattc aatttggcat tggatactct tttggcttca cctgcagaga 214320  
tttgcgatct ggtgacgagg gaaatggtaa agaccagtaa cgatattgtg cataatattg 214380  
gatcgaattc caacacggac gcgcttcaaa agagccttca agttgggtgcc tcagcagtag 214440  
aaaaatcaga cgagtctact ctttctacta aagaaactga cgtatattcc cttgtttctg 214500  
ctctggctaa gagcaaatct cctctatctt cttcctcatc tttgtcgtct gagggacatc 214560  
tcacctctaa ggagattgat aggacatgga acacccccgc tcttctcggg accgctaaaa 214620  
ctacatccta ttctgtttct gaagacgctc tcaatgctcc tctttcagcc gtgttggaact 214680  
ttagaaggaa tgttgtggat gctactaaat ctctgtacga agttgcagct gtttgtagt 214740  
tcatgagtaa agaggaggat gtgcgttctt cgagtagaaa gattatggga atgggtggaac 214800  
aagaatcgcc cgttatgcaa gacattggca ttgaccgcat tgctagtctt gttagtacag 214860  
ttgctacccc caaacagcat cgcagattct tacagacagt aaacgattac aaaaattatc 214920  
tcattagaaa agttgcatcg aatccccctc tctcttcaag attgggagga atatccccta 214980  
ctagtgttaa caccgattac aaccttaaa ctgtatatga tgggtgtgtt tcttctcat 215040  
catcaatgac cccctcgtcc atgtctgtct ctgacagatt ctggtcggga gtattttctc 215100  
agtgcctaga gactggccct tcaatgtttg ccgatgctgg tcatggaggt agtaacctgt 215160  
tccaaatcac tgcacctaaa ctttacgggt ctagagtcaa cacctacgca gctctgagct 215220  
ctggcgttga gcggctgaga gactctattt cttctgcgac tcaggaaaga aagaatagga 215280  
ttgcaaagag catcgaagct ctggaacgt tcgtaaccga tgtggtggg ggagatactt 215340  
tggatcaatt gcgtaaggcc cagaacatgt acaacaaact gtcagatatt acttccaact 215400  
ctatctatag tgatttcgga aacattgact gcgctaaaat catgaagaat gtgacagca 215460  
agaaaatgac cgctagacaa caatcagata ctattcttag ctctctcttg cacgaactcg 215520  
ctggcctggt acacaaacaa caacctcaat tggctactca atttgctctg gcgagccatg 215580  
ttatcaaggc aaagtatgtc actaatgacc tcaataatat ccacgagaag gaaacattca 215640  
gtcaattgat ggccgtggcc ggtgttgccg attactataa tgtgtcggca gctgccatgt 215700  
gtcagcgtct agttgtctcc gacgtaacaa tgttctcgg cggaaccatg ctccaacaag 215760  
gcctgttctg ttcatctctt cttaacacag tactttctc ccagggttct gataatatta 215820  
aaatgaacga attgaacgat gaaacaaagt ctcttttgg taaactggta ggattttgctg 215880  
gtacagtttc agatgcgcta ggatctaggc acgtgtcttc aattagacgt gtacagaacg 215940  
aagaggataa gaaattagac aggagttttg ttacatcact ttattcagca tacagagatt 216000  
tgaggaagaa gactgaacta tacagggaaa ctgatactat taacaaactt ttcggacatc 216060  
aaaactttat gtcttacgaa tcttccatgc tcaagaggac ttccttggtg catgacgtg 216120  
tttccggccc taggccaaga aggtacagca cccttgagga tgtacttgag gctccttcca 216180  
cggttcacaa atcgttcatg gtttcttacc cagagagggc agctgcttct aggcgagtga 216240  
agagggctgg actcagggt ctggctgata acaggatgga atctctttac ggggaagaag 216300  
tcttgaacga tatgaggtct tcggcgggtc cttccgaaat gatggatata gagtatgggtg 216360  
agggagatt catgatgat attagtatga ataggatga tattgccttt attgtattccg 216420  
aagaagagtc tgaatcatct actgatttct cctcatcaga tgaatattcc gattcatccg 216480  
atgagtatga ttttgatgat gataataatg gccagtctcc ttattcaact acatcttatt 216540  
cgtatgatgc tctagaccgt ctgaattctg ccgctaagcc tcttactgcc atctacgggt 216600  
gcaggggaga aggtgaagac gatgaggaaa atgacctcta tgaagaagaa caagaaagga 216660  
ggagacgtcg ctcatacaag atgggaaga tccttagaga tcttcatgag agtgatgatg 216720  
acgacgatga ctactttgat gacgaatttg atggcgacg ttcaatgtca gaaactattg 216780  
caaccagaag agctggccgt attcaatatg gtccagggtt cctatctcat tctaatactc 216840  
ttaaccgtcc ggctaaagca cgcgtttct tgacacgagg caagaaattc aggccttctg 216900  
cgtacgatag attctttatg gaggatgacg attccctcct cttctctgac gaatctacca 216960  
cttcttcttc ctcttccgat tctccattct cctccttcag caaggggaga aaatgcaagc 217020  
gccgaacaag cgaggacca tgtgcctttg ttaagagagt tgtacgtgct tttgtgccca 217080  
ccagagtaac aatgatcaat ggtcgagtta gcatgatcac cccagtgact agtgaaaata 217140  
cagtaggatt ctatgaaaat taccagaagg ccaacaagag ggaaagggcg cgtctgatcg 217200  
aagaatacaa aattgttaag ggtgcttcag ctaccttgcc cgacgaatac gtagagggtg 217260  
gagcatctaa acaagtgtc cccagggaac ttagggaggtc tcttatcaag gcagctgctt 217320  
atgttggccg caccacaaga agtaacttga atattatctt tgacgctctc accacaacat 217380  
caaacgccac tctagttaat gacctctcta ctcttttggg tgatacactt ttgttcgcca 217440  
aacaactaga ggcaattacc gagaggagga ataggcta at gaaagaccta actgaaatct 217500  
ctccttcaact ttccacatca ttcggtgatg caagtaagaa caccacaaatg atggccgatg 217560  
ccaaacagat cgtttcagga ggaaatttca agtctgcgg ctatctaggt gtccctctca 217620  
gaactcttgc tctcatgtatt aagggcacta atacagttga tcgtcttttg gctaccaaaa 217680  
ataagaacca tctcgaatgg atgaccacag ccgctattgt ttttgcacgt tcattcaacg 217740  
atactacttt ccatgcactc gaagatacac taaaaatgac ctccgctttg acagacatgt 217800  
acagcgcttt caccaacctc gtcggatcgg aacattctca gcgtctaaaa gtaaagagta 217860



ctctttttaga	ttctattttct	aacactagga	tggtctcacac	tgaagcagtc	atgggtctcg	217920
tataccctac	agcgttcac	aaccatgaaa	tgccctctga	ttacacacag	cgcagagaga	217980
tgcaatcact	cgctcttaac	attcttaggg	gagtttaattg	tagccaattg	ccacgaaagg	218040
atattggaga	cactgctggc	ctggtgacct	ttattacatc	acgtaaattt	gcagggttatg	218100
gaggagaaa	gggaggtttg	tctctataca	gaatgtccat	tggtgatgct	ctttcttgcc	218160
cctctgacaa	tcggctcaag	ggagcagtc	ctctagaggt	aggaaagtgg	caggatatgg	218220
gagaggaaat	cttctacaag	aggagcaacg	atctggtcga	tttttgttca	aagaacaata	218280
tttctctgga	aaatgccgta	ggtcctattg	ctagggttgt	tcccaatgga	actaacatgg	218340
ctgatattgg	catgaccgat	atcattttcta	gaacagtcaa	ggatgacgct	tcaatgatca	218400
ggcttaggcg	cgcagaagag	ggcgtggcg	cagcaggaaa	attcattaca	gcctcagcca	218460
tgggtaattt	gtacggagg	attgataccg	ttgtgaacct	aactgaaaaa	ctatacgact	218520
cggtcgttct	gctccaagat	tcagactcgt	tcaatacacc	aacagaaatg	gccactgcta	218580
ttatcaaccg	tatgaagtcg	aggaaacata	aggctctcaa	aacaccattc	gggggagata	218640
ttgccaccta	taagaacttc	ccatcctcct	ctgaagcaat	tgtagttaga	gccaaaggaaa	218700
tgcgtaactc	tattagcact	atcgtgatgg	acatttccaa	gtcaagggga	atcaactcgt	218760
tctcatcccg	cagtggttct	actttggcca	agatttccac	atctgaattt	gaaaggatac	218820
tagaaacatc	tgctgttctt	tcaaatacaa	aggccaatct	gagaaccatt	gagaatagac	218880
tcgccgaaca	ctacaacaaa	ctaaaacaat	tcagccatat	tagtaatgat	ggactttccg	218940
agacacgcgc	agtcgttgcc	gtaattgctg	aatctttaac	ccccgtgtat	gcggatgaca	219000
ccagcgagag	aggagcatct	gttagtgaac	tattgacaga	caatactctc	ctcaaattta	219060
ttgttcaaaa	tgaactgaaa	aacattgaag	aggcaaaacg	tcacgtgacc	gccgcaattg	219120
aaggttcac	ccaactgcac	gaaaaaatgt	tgagcctgct	cgttgcctca	gccgacatca	219180
accgtatgtc	cgcccaaat	aacctcgaat	gtaagaaatt	gactgaagga	aatagtaact	219240
ttgtaccaat	gactaacgac	caaggtggta	catcataaaa	gcacaaagaa	acagggtatct	219300
ggctgaagac	cgatgaagaa	aataacacca	gttctatcaa	ggacaatgat	cagcgtagag	219360
tagctaaaa	catcctcgca	attgtagagg	acaatagaaa	tgcaaccatc	cgttctcgct	219420
tacagtctct	ttgctttgga	aaatatgcca	tgaacgacat	ttttgcactt	gatgatgcgc	219480
atattaagaa	tatggacaaa	ctcattgaaa	aactaggcga	agcactcgca	gagaaggcat	219540
ctccttctag	ctcggccatt	tcttcttcc	catcatctaa	cacaacatcc	tcctcttctt	219600
ctcccagttc	ttccccatca	tcacatctct	cctctttctc	aatggactat	tcaaacaacc	219660
ttgccaaaa	tatccccctac	atgcctatcg	tcttccaaaa	caaacaatct	aatgtcaatt	219720
cttctcagcg	atcactctca	tcacatctat	catcttcttc	ttcatctgcc	aatattgata	219780
atgttgagca	caaaaaagtg	gctctccaac	aacttcaaac	acaagaatct	aacgatttga	219840
gtaacgtact	ttctgttacc	accaagcaca	gatttgcgtc	tcataatcaa	gctgcaactg	219900
ttggcatctt	caacggaagg	caacacgcag	agacagttgt	tgctatacca	aatgcaaaaca	219960
aggctaataa	taatgccacc	gtttccgcag	gccaaaggaat	tcttaccgcg	ttctcagccc	220020
ctgaaaatgt	ttcctccacc	agcatgcaat	tgcttccatc	atcatcatca	tcatacaaatg	220080
gagatgataa	taaggtagca	gtaactgtca	ggcttaacca	gtacgccaac	tcaatcttat	220140
catctattga	aaacgcacat	gaattttaagg	acttgaagga	agcagaaagg	aaaatcgatc	220200
tggccatcca	ggcagcttcc	accacagaaa	caaaggaaat	ggtcaccgtg	tctaagtgcc	220260
cctctgctaa	ccagactgcc	atcactgcca	tctctcaagc	taaatccctt	aagaaaagtg	220320
ccctcgaaat	attggaaaaga	gttatcaagg	cagtcgaggt	ttacacccca	gattcatcta	220380
ttgcagccgt	ttctcttccc	gttaatggag	attctatggt	ttcttcttcc	tcgggatcag	220440
gatctgctcc	ctcttcatca	tcctctctct	catcctcatc	ctcttctctt	aatgtgacag	220500
actatttcaa	ctatgcttac	ggaaaattga	agaacattga	tgaaaatact	gaagaagggg	220560
cagaaaactgt	ccagaaaaaac	atggtcgaac	aagatgctgc	cgttcgcac	cctcttctag	220620
tatcatatgc	tccattcagc	gaaatgatga	gcgctgctat	tgacaagttg	aacgaatact	220680
accaactgat	tgatgccatc	aaaacaaaaga	tcgtgtcaga	cactaaacag	gcttctctat	220740
gggccatcaa	ggaacaggac	aaggagcttg	atatggacaa	agaacagggtg	atttcaaaga	220800
ttaataactt	gcaacaaaac	ttttcaaacg	aatcagacaa	gataaagatg	gctatttagtg	220860
ttttggacaa	caaaaaggaa	gaatttagagc	ttcagaacaa	caaaactagg	agcttttattg	220920
aaactacaaa	gagccgtatc	gaggctgggtg	gaggagatgt	agcaaaactc	aaggagatta	220980
tcgattacga	aaacacatct	gaaaatgaca	acaactctct	ccagagcctg	aaagcattcg	221040
ctgctgataa	ctcgggggaca	gtttacaccc	ccactgacat	gagcaatgga	agagacacaa	221100
aatcagacag	taaatttgct	gacatgtaca	acaaacagat	tctcgaggga	ggaatcaaac	221160
tcatcaatga	gggacaaaat	actgtaaagg	tagacttttc	aaaggctttg	gaggctttcc	221220
ctagacaatc	caacgggtgt	tcagagcctg	tatcttcttc	agttgtggag	aggagacagc	221280
gagaacgtct	tcaggctgtc	gagatgttta	tgccaataat	gatggagcgc	accgagagtc	221340
tgaggaagag	gttggcgat	tcggctgtct	agtggaaatac	tgtaataaat	gtagaagaaa	221400
ctgttaatag	tggtatgggt	aacatcaaga	gtgaaaggct	cacagagatt	aggaatcaag	221460
cacaaatcgc	tgaagcact	gcactaaact	ccatcaacga	cgagattgta	gagtctcctc	221520

tcaccctctc	tttgggagca	cgagtcgacc	agctcttgat	caaggtagat	agagtaggaa	221580
gtatccaaca	gcagcaacag	caacagcagc	aacagcagca	gcttcccaaa	ttgacagcta	221640
cagaacagag	aaaggaacaa	caatacgctg	cagatagggt	tgtttacgat	ccttcataca	221700
cctgcttcct	gcaacctctt	cacgagacaa	ttaaactgat	ttctctctgc	tataattcaa	221760
agaacaagg	tcctctcagt	aacacacgtg	gtgttcccac	tagcgatgcc	gacttgcac	221820
tgatgacat	cactgacttg	tctaggtctg	tactcgactc	ttcttccact	tcctccaaga	221880
aaatgctgta	cgaaaatggt	ccctcatcaa	ttgttctctg	actctgccag	caatgcgcaa	221940
tgatgatcac	caacgtccac	gaagccactc	atacttctcc	tcattcattc	aatttcgaga	222000
acaaaagatc	cctgaagcag	ctgacagaaa	tggtgaacgc	tgccacttca	tcagtgacg	222060
gtcctgccgt	gagacacgat	gtactaacia	tggttagagtc	caacaatggt	taagtcaaag	222120
attttggtt	cactcaccgc	caaaaaggtg	cgtgtatcac	ccctgttaat	acacttctgg	222180
gaggtacttt	cagtggaaat	gttgacacta	atactgttat	ccttccctact	tctgagttgt	222240
ttaactgcc	aggagttgaa	aatgacaaat	ttagatccat	ggttaacagg	acaaccgaca	222300
agaatgtggc	tgacgcaccc	aagtcatctg	caagcatcgt	ggagactcct	gctgcacagt	222360
tcaccaacgc	cgagcacctt	tacttcccct	tcaaggacca	gaggcgacac	ttcaactcca	222420
ctaccgacgc	catcatttct	ggatgagcgc	gcgaatcttc	atctcaattg	aacactactt	222480
gtgatcaaaa	tctggtaaac	attgatcaaaa	ctactggcct	cccagtggtt	acaggaagaa	222540
agcaggcgca	aagaaggatt	gtgcacactg	aaaacactat	ggaaggagct	cgcaaggaca	222600
agaacagtgg	catcccttca	tgtacaaagg	accgtcaaac	ttatatcgat	atgggcacaa	222660
aattcatggt	tgctccaggc	tctcttctga	atgctaacia	ggaagaaact	ctccgtctaa	222720
acaggctttc	agacattaac	aacgtgagac	attatggcac	tgatgttcat	gtggcaggcg	222780
caaactctgc	atggagaatt	ggtgaggtgg	tgagagccgc	ctcctcattc	cctgacggag	222840
ataaggaaac	ggctatgaaa	aagatgcttc	ttctaggatc	tgtatctgcc	atctctgtct	222900
aaaaatctgc	cagtcacatt	aacgatccta	ctgctttggt	gagcaccac	acgtctatcc	222960
agaatctggt	caaggaagct	ttcccagacc	ctggttgttc	ctctaattac	ttggggctctg	223020
ctgaatctac	gttcgccact	caactcgctt	accgcagcg	cctgttccct	aacggagatg	223080
acgaaaacgt	tacaaccgtt	tcaaatatct	gccctatgga	tttgatggga	agtacaaagc	223140
gctacaatga	cgctttcaac	aacatctttg	gctctaaaat	gacatctact	aacaaaaagg	223200
ggtcaaattg	tgaaaatcta	ctgaaatctg	ccatgtctaa	cgttcctgct	atcaacactg	223260
cctttggagc	ctttgaagaa	gcttcatctt	ctgtcaggaa	taggctttct	cccctttatg	223320
aagacagcac	caaatattcg	tccaaccaac	ttgctgtaca	ggccatgacc	gatactgctg	223380
tgatgcttt	gtccgctgtt	tctactgttg	tcggtcgcca	gaatggcaga	aatactcttc	223440
ttctctctcc	tacttctatt	acttctatcg	caaccagtgg	ccgtccatca	ctctcttatt	223500
cttcggacat	gaaatctaac	ctcatcaaga	caatttcccg	catcaataga	gacgctagcc	223560
tcctgtcaat	gggagacagc	caagtactcg	caggttcttc	cttctttaac	tctttccttc	223620
gttcttcttc	catccctgtc	accaccagcc	aggatggaaa	tggtgcagca	gcggaaattg	223680
ttctggggac	tattctcgac	aagactgtgg	agatcaacia	gagattcgag	atgcttggag	223740
gaggaaaaat	ggtcgccggg	agtctgaag	ctcgtgccat	ccagcgcaat	acaatgtcct	223800
ctattctcca	gatgaacgaa	aatgaactcg	ctcgtgactt	gtgcgaaatt	gaaaataaaa	223860
ttgagactag	gcaactgagg	gatgcttcc	aggatttgaa	gaggtctatg	ctgatgactc	223920
caggaggcgt	gggagccatt	tcttctggag	caagtaccaa	caatgttccc	ctttctcttc	223980
tcatgtcacg	gtcgtatgca	tccagcggtc	tctgtatgaa	caacaacagt	gccaatgtaa	224040
tggaagctgt	ggatagtttc	aatactactc	ctttgctcgt	taggcacatg	atgttggata	224100
gtggaagatc	ccccgttccc	atggccaagg	aaattaggag	catgctgacc	caaccaagag	224160
ctctcaccgc	ccgcgctcta	ctgagcgaat	cttcccctct	tctcactgaa	atctgcctct	224220
acaacacccg	cgacactcaa	ccagaaaggg	cagtcgacag	actactaact	tcagcctatc	224280
tagtaaaaaca	agctaaaaga	ttcgacggag	ttgacccagc	cttccctgcc	gccctcacct	224340
gcgcttctca	cctcatgctt	tcttccatgg	atcccatac	aaagtcatct	ttcatggaca	224400
acatcaaatt	gcacatgact	gatactcaat	gcttcttcaa	gaacattgaa	cgatttgaga	224460
aattcttggg	aagatatggg	gacgaatacg	ccatgtccca	caagcaaaat	tgtaactgcc	224520
ccttccatct	ccaccacact	tttactccct	cagataacga	gcactctggt	tcctcttttcg	224580
cattcgcccg	cccagaagtc	tccatggaag	aaatttagag	cacaccctat	caggccaaca	224640
agcttattag	tgacaacacat	tacgtgatga	acatgtccaa	gatcgattct	agagtaaacag	224700
gatcttccct	ccttaagaag	gttagcgaat	ggactgaaat	gagaatgaac	tccaacttta	224760
atggaacatt	tgaaccatca	agactcgccc	tctccaactc	tgccatgaca	acggcaggag	224820
tcaacctcga	cggtattgtc	aaaccaaata	atgcaagaag	tgtactagga	atattggaat	224880
gtcatcgcca	gcacgtgtgc	accgcccagc	ccaagggaac	gtcgtcttca	gccatgccag	224940
ccgtcttcca	ggcaaccgat	ggaaacggt	acgaatctga	actgatccag	aatgctctgc	225000
caaggaacag	atacatccaa	aagagcacaa	tgaacgctca	aactgtcgtg	tttgctaatg	225060
ttttggaaca	acttatcgcc	gatcttggaa	aggttatcgt	gaacgaactg	gccggcacca	225120
tcgctgaatc	tgtaccagaa	agcgtatatg	aaaacaccaa	ggaaatgatt	gatagactag	225180

gctctgacga cctcttcaaa tctaataata atggaggagtg agaatcaatg gattatgaag 225240  
 atagcgaaac aacatccaac aatgggtccc tcctcatctc agaagccatg aagaatgccg 225300  
 tctatcacac actaatttcc ggcaaggcag ctgcgccgga aaatgtacca ttcgcctcat 225360  
 gcgccagcgg ccctctcgcc tttgatttcc ttctgtcaaa gggagataca ttcgaagaaa 225420  
 agaacgcgca acaagggtgca gcagctgccg tatcctctac ctattcttcc tcttctaaca 225480  
 ctactcttcg taagcatttg gctcgagttt tcgaagccat ctctaagcaa gtaactgatg 225540  
 ctgaattcaa ggatattctc aacgatattc aacgtaatat ttcttctgac tatactaact 225600  
 gtccaccaaa tactaaccac aatgcctttg ctctagctat caagagagaa ttcagcagaa 225660  
 ttgtttcctt ctttaaccatt ctctgtaaga acattacacc cgcattagtc gaccctaaag 225720  
 gcgcgttaca cgagaaagta gccatctatt tgacccttct ttcaaccaaa tcaaaactag 225780  
 aaaacttttt ccaatacggg ctacagtaatt cgtcctcagt tgatcttagc catctaaaac 225840  
 ccattaattg tagcaacaat gtcaagaata ttgaagacac attcatgtac agaaatgtcc 225900  
 accctattct tattatggcc ctcccagaaa atttcacagc tctcttgcaa caggaacaaa 225960  
 tggaccccgga tactgccatt gaaagcagac gctcccttac caccttccct aatcatccca 226020  
 acatgtcttc aatggcgaaac ggtgcaagag ccgctgtggg tgcaggagga ggaaacccaa 226080  
 tgggcttgta tctttcttcc cacatttctc acgagctctac cgtcacaaca tcaaaccocg 226140  
 tcacagacac cacagaaaac gtcaactatc attcctctgt tacacaagat cctgttatgg 226200  
 tagtgaaccc cttcaaggat tctgctaggt tgatcggtta caacaacaat actggaattg 226260  
 atgtcttgaa tgataagtgc tgcaactact tgcaagtatc catgccatct gaatcatctg 226320  
 gcctcgtcac caatactgga tgctcttctt ctcttctctc atcttctgtc gataccttca 226380  
 agtacctcag gagagacaat acgcctgtga atcttcccc tgctcacaca gccgttctct 226440  
 gttctgatgc ttctctaat ctcttgagcg tgttctccag ggcagatatt gtcctcgaaa 226500  
 acatgaacgt gagatttggg ttcattgccg agattattgc tgccgtctcc aaattcaagg 226560  
 ggctgaccaa ggaagagggt attaaagcaa tggtttctca gaacaacatc aacaacaaca 226620  
 gcaacaacaa caacggaaat gggaagaaaa caaccgtcga tccagtcact ggggatattg 226680  
 ttatcaccaa tgccacattc cccgacactc gtctctata cactgcagca aatggaggaa 226740  
 catcatcatt caaatgggga gatattcaacg acagaaaaat gcacgccaag gctttcccca 226800  
 ccttctttat tggtaaccca accgcgcgcg caacagctaa cggagtgcct cttacatctg 226860  
 agggaaattc cctcactgaa gaaaaacgca agaaaatcgc aggcattctc gaaggatcaa 226920  
 ttggcacggg ggctctgcgt gcagccgcga acaccgcct ctcatccgac atggaacctg 226980  
 tcatgaaggg atggaacaac attgttcagc ttcaacaaac attcaagaaa gcttcagata 227040  
 aactcactca tcttttgaga tcgggaggaa ttccaccag aagccaagaa acaaaccgta 227100  
 ttattaacaa gatgcacgac agcttcaaga cattggaggga atgtcgtagg gtgatccaag 227160  
 acgaggctgc tctgctcgtt gccaccagcg atcttttgac cgggtggctac ggaggagatg 227220  
 ctgctctggc catggtttct ccagtagctc cagaaatgac tggctctatt ggtgcaatct 227280  
 ccgcgccagt tagaggtatt agccacttgt tgaactggg aggtgtttct gctgctaacg 227340  
 cagctatccg caagcgctc aacctacct catccaacgg gaaaacacta ccagaacatg 227400  
 gaatcgtaca caaatcagcc aagacacttt tgcttgattc agactctatt agcaacctat 227460  
 acaacactga tcttcaagac gttgtctcta acgctaggga taacaacaat ttgggaagaa 227520  
 ttatgcaatc tttgggactt aaggggaata atgcagggga tttggtttat tctgctagac 227580  
 aactgacgga ccttattact gtaccagaat atggaacaaa tcgcgatctt accaagcgtc 227640  
 aagctatcct taaaatgctc atttctaacc ctgaaattct agaaaatggt gcagatacca 227700  
 ttaccttac aacaggtaaa aatgctctcg caccggatat tgctcaggaa atggcttgtg 227760  
 cgtctctgac agtcggaggga agtgaggagg gaaaactgtc atcagacgac aatgttcaat 227820  
 ccctgaaccg cctttatttt cgggtctaga ctagtgtaat attgtatggt taattttattg 227880  
 tgtaacaatt ggtttatatt tgtatgtgat ttttctggac aaataaaaaga attggaataa 227940  
 aagactttgt attttttacc atttttttta aacacctttt tatgccacgg 228000  
 ggttgattgt agcatttctc cccctggag atttcatagt gggacatcca aagagtgtaa 228060  
 ctctagtgc tgatacccc aaacctagga ggaatccggt aagtgcagga gtaacagaaa 228120  
 gcatgaaagc gaccatgctc aatgcgatga cgacaacggc caagccgacg atgatgtttc 228180  
 tggatcatatc agacatgatt tctggtatct gtttataggg tttcaccag aagtgtttta 228240  
 aacccaaaag tataaatttt attctatact gctcatctca ctcttccgac ctcaaaaagg 228300  
 agaaaagagt cagctcttgc ccagccacca cgcaggcagg agctcctcgg acgctcctgt 228360  
 atttttgtcc agaaatggat aaagtttgtg ttatatcaaa tactaggagg cgcacgttta 228420  
 aggtacctgc cgatctactg tgcgttgcaa cagaaccgga aatctctacc aaggaagaag 228480  
 atgcaggatc agaaatcgag acaagagtgg tgggtgtttc aagggtcgctc tcggtccagg 228540  
 aactacatac aataaatcca aatgatgaag gatatttctgt ccaacttttc aaggactacc 228600  
 tgaatttgca atctgcacaa ggaaaaaaac ccatgtgttt gtacatccaa ataaaggctg 228660  
 gagaggatct tgaaggaga ttaatcagtg gaggaactgc atacctggat ccggcaacac 228720  
 accttttcta tcttgatttc tccctttacc ctaattattc aatattcaat gacatttcat 228780  
 cccgtctgaa aatcattgat gaagacacgt acaatggtgt tgttttctct aacagtgaag 228840

aaaaagagaa	ggatgcacta	gtgctgataa	gggtgacttt	ttctacgcat	gaaaaggcaa	228900
ttgaagcagc	cataaaaaaa	ataatgctaa	ggaaagtgtt	tttcaaggat	ggagatcttg	228960
atttcgggta	cttacgtata	ccaaaatcta	aactggacaa	atttactccc	tatttttcgga	229020
gtcaatacgg	gatagtaaat	gttgaaaaaa	atatccctgg	ttacatatgg	ggagaaatta	229080
tgaagcaacg	agtgcgatgt	tccagatggt	acctttacaa	caccgactcg	gaatgggaat	229140
ataaaaaatgt	ggccgaagaa	agagttggac	ctcgccagtt	agtgaaaaaa	tatggtgcca	229200
agtgtgaaaa	tttatgtttt	agggacatag	acctcagaaa	aaaggaagca	aaggaaaaaa	229260
gggatataga	aagagaaact	gaaagcagat	atgtggctgt	aacactaacc	cataagcatg	229320
aaatgcctga	aaatatgccc	tatttttgac	caaagtgttc	agtggtgagg	ttggatgaaa	229380
ctagaatact	tttatgtttt	gtggatgaaa	tttcttataa	tgatgaagat	gtagacgaaa	229440
ttttgtctga	gaatgatca	ctaagaaatg	tttctattag	acataaggaa	aatgtacctg	229500
tacacacggt	attaaaaaaa	ggtgtgtcta	ttcatgctag	atttaccctt	aatggtttgg	229560
atgatgcttt	aataatttta	aagagaatac	caaaaactta	ttttgaagat	gaggaactac	229620
aagccgcttg	tgcgcgatgt	aaccttgaac	agtacgaatg	gctttgttct	aataatagag	229680
ggaataaagt	agaacatgta	aagtcgcggg	tagtgactcg	agcagttaag	cgtaggagaa	229740
aatgtataga	ctggattttat	tttgataaa	acacttttaa	tttaaaactac	aaatactttg	229800
ataaaaaagt	tactgctagt	atggcatcta	aaatatgtaa	tgcaaaacac	gactgttttag	229860
ttttccatag	aaaaatggaa	ttggaagatt	tgactgagag	cgcatatttc	aaggtagaac	229920
cttccccaat	aaattttgcc	aagttaaaat	cttgcccggg	tgtaaataat	gtgcagaaaa	229980
aaacagatgg	gcatttttct	gttataagat	tctttagaaa	catgacaaag	ggtgatctta	230040
ttcaaaaggat	ggatcttttt	tgtaggttta	ttcccgactc	acacactatt	acacttttga	230100
gtagggcgga	tttttatgca	tgtaaaagag	gagaatctat	gcataatgtc	acaaacaaac	230160
accgtattct	tactacaaa	ttctccaacg	ctccccatgc	ggccatcgaa	caaataacca	230220
atatcatcag	tgatacaaga	ggacgtaagg	gtatacacat	agaatacgcg	atcgaaaatg	230280
tacaagaaat	gtacgaagaa	gatggaagaa	gatatagaag	taaatacact	ggaactttta	230340
ccgagtacaa	aagaaatgag	gacaaaacct	tcaaattctct	tcttgctcct	catttaacac	230400
ctgtcaacaa	accatataat	attaaccatt	tgtatgagca	atatggaaat	tttgatgaag	230460
aattagaaga	caagttgagg	agtggtttca	tttcttatga	cacgtatgtt	actgcaaaag	230520
acaactgggg	caggtgtgca	actggaaaag	gggcgtgcat	ctaggaccca	taactcaaa	230580
tcaatactgg	tatgtgtttt	ttgagaaggt	gtcagaatgt	gcacattaaa	aacatacaaa	230640
atgactactt	caacagaaat	atcaaagaac	ctttcagatg	tgttatccat	caaggcaact	230700
ggagattggt	gcagtaatat	caagacggta	ttttcacctt	tcacagaagg	caagggaat	230760
ttaccaaaaca	gtctcccgtt	tacgagaagt	cccaatacaa	catgtgggtc	aagagaggcg	230820
gcaaacgcca	cagagcatth	tatcacccgc	tttgcaaagg	acaaatatga	gcggaaaaga	230880
gtaaaacgta	caatcggatt	caccctcgac	aacacaaaag	agttgacgcc	caacagatac	230940
ttggtagcag	atgtatactc	ttggcaagaa	gagaaaaatg	tgtttgaaag	atthttgtgtc	231000
ccaccaggaa	agtcgggaac	atthgtacgc	tactctaattg	aagataaaaag	ttttctacta	231060
gcagataccg	gaagatatat	gaaaaagaag	tacgatgatc	cagaaaaata	gaccagtagt	231120
gggggtgatg	atgacgatga	cgacgatgat	gatgatgacg	acaacaacaa	tgttgacgtg	231180
tatgaagaaa	acgacccag	aaatgtattc	gaggtcgaaa	aggatgaaaa	atatgcctgt	231240
actthttcaa	ttthgttcta	tagagcaatg	aaaaagtcct	ctcctgtatg	tagagggtta	231300
ttagtataga	cagatggacc	ctcatctcac	ccataacggg	ccccgtcagc	atthaatcca	231360
ttcggaggaa	gttctatgtt	gaacggttat	ggtgcagggtg	cagatgcact	agaagaagag	231420
gatgaagttg	atggagttcc	tgaaagagag	aggattacaa	atthttgtct	caagagagga	231480
cctgcaactg	gccagaactt	tgtatctgtt	aaactggaac	atgatggatc	taaagcagac	231540
ctgtacaacg	tcacgtgctt	ctccaagcag	cgtggagtat	aaaaggcggg	acgcatacgc	231600
aaatgcaact	cattctttct	catcatctaa	ccatggctgg	tcgtgtagag	ctcgtcactg	231660
gacccatgtt	tgcgggcaag	tctacctacc	tgaaaaacat	ataccaacaa	gaaaaatggag	231720
gcaataaaca	ttgcctgttt	gtcaaacact	ccctagaaac	taggtacggt	tgtggaactg	231780
gaacaatagt	cactcatgcc	ggagaagtga	ttgaagggtt	tactacagtt	tcttctatca	231840
aggaactaat	cagtgtgtta	ccagaagtgt	tggtatgtgt	tctcattgac	gaagggcaat	231900
tcttcacgga	tttgggtgta	gtcaatagac	tggtcgacaa	ggggaaaagg	attgtgattg	231960
cagcacttga	tggaacttct	gaccagcaaa	tgttcagctc	tattcataag	ctattgtcctt	232020
atacaaattc	cattgttaag	ctagcatcta	aatgtatgat	ttgtaaaatt	gataccaaag	232080
aagctccttt	tactgttaag	tttggtaatg	acaatgataa	taatgttata	tgtgtaggag	232140
gagctgaaat	gtacgctgct	gcctgccggg	actgttacaa	aaaaattaac	aagaaaaaga	232200
acaaggggaa	acttgttgta	cttgaaggag	gtgacagggt	cggtaaagag	acccaagcca	232260
aactcttggt	gaccaataaa	aactgcctc	tttatggagg	agaatatatg	tgctttcccg	232320
acaggagcag	ccatacgggt	aaactcatca	atgattatth	aactaagaaa	attgaactag	232380
atgatcatgc	agctcacttg	ttatthttctg	caaatagatg	ggaagtttgt	agtaaaatta	232440
agcagttggt	agacgatgga	atccatgttg	tgtatggatg	atattactac	tcggggattg	232500

ttttctcttt	agctagagga	gtggataccg	ttgagtgggt	ctctgctagc	gatgagggac	232560
ttcctcagcc	cgatcttgta	ttgttgatgc	tttttagatgt	tgaaaagtgt	tcaaataggg	232620
atactttttg	tgtcgaaaga	tttgagacaa	attccattca	agaacgtgct	agagccctat	232680
tcctagacct	cgcaaaataag	gacgaaaaga	atgtatggat	taaggtagac	gctcgcggca	232740
ccatttgagga	ggtgcaaact	aaaattataa	atattgtata	taatatgtgt	gaagaataaa	232800
gaattgtaaa	ctgtttccat	gatgtgtttc	tttcctagt	attttcacat	gttacctaga	232860
aagactttgc	ccgacactga	aaatggttat	tttgtcttgg	acgagtctct	tctggagaag	232920
gtgtactatg	ataacaacaa	tgaactgatt	gtaagagttg	gtgggattta	tatgcagata	232980
tgcaagtcaa	aatacatctt	ccatcacgat	gatccagaga	ggttctttta	tagtgtgttg	233040
gaggattatc	accccatcaa	agagattgtt	gaacgactag	cagaagagga	tggggatttt	233100
ttaggaccgt	gggagttttt	atcgcgcaaa	caagtgaacc	tccaacacgg	gtgctacaaa	233160
gctcttttgt	cattgccaga	ggacaaatat	tgtaacctat	tattacccca	gcaaatgaaa	233220
accaacctgg	aaaaaatgga	agaaatacag	cgtactagac	tcattcactc	tagaacgtac	233280
aatacacccc	agatagaatt	gtctgaccag	ctagatggat	gtgttatatg	ttaacacttt	233340
ttcagtagca	tgtaatat	ttgatatcaa	attgaaagga	aaatggttca	gaaattggta	233400
gtaaaaactt	tattgtttct	ttcctctttc	ttttgtctgt	cactttattc	gtcagacata	233460
acaaccataa	tactgagtat	gggtcttgga	tctttgttgg	gacatttttg	gtgtgttaaa	233520
agtatcacag	gaaaaatcgt	ctcaagaaag	gcaatggata	attataggct	cgtgtcggtc	233580
ctcggctctg	tatgtggagc	aggaattttc	tgtggatacc	agcaatatcc	tacacctcag	233640
tctgcaccca	tggcgttgct	gtctgtcttc	tgtgtcactg	ctactgctgc	tacatcaact	233700
attgtaggat	tctggatata	gatgattatc	caggcattga	aaaggagtaa	taaagatata	233760
atataaaatt	tgtgtgtttt	atttgtttga	taatacaatt	ttcaccttgt	accgaaaccc	233820
attctctaca	agcactgtgt	ctttggacac	ggggatcaat	atcttggatt	ttggatgaat	233880
gttcctgacc	ctcatatctt	cttcattttc	aaatacgagg	agggatgttt	tctcagtatt	233940
cctgttgtgt	atttggaggt	gaacagggaa	ggagaatagt	tgattgttct	ttggtgttag	234000
atggaactcg	tataccatga	attgtactcc	tccaatatct	tcaagataga	tgcattgtgt	234060
atttttgtca	taccatctcg	tatggctcgc	agtgtcagg	tacaatctct	tgtgtgtgtg	234120
gttacacatc	tttccactgg	tgggagagga	agtattatgg	ttcaatattg	atggaggatg	234180
gattttttata	ccctcatgcg	aatgcgatgt	ggaaaatttt	atatgtttcc	gacaagggga	234240
aaaatatgaa	taaactgagc	atacacatgt	cgtatccaga	tatctagggt	gggaagaatc	234300
cagaaacgat	tcaaaccatt	tctggaacat	tcattttctg	aaagggttac	attttattgt	234360
atctggtgca	tattttctgg	taccattttt	ggcacctctg	tgtaatctga	cattgggtcg	234420
accagaggt	ccaccctcgc	aacttgacat	taggccgacc	cagcgggtcca	ccctctaaac	234480
togagtgagc	tgaaaaaatt	tttataaaat	ttttgatgag	gaaaaataga	agtaagatgc	234540
ttgtcaatgc	atgctccctg	tggaaagtgg	tggatgagat	aggagggaga	ggctatccag	234600
atatctagggt	cggaagaat	ccagaaacga	ttctagccat	ttctggaaca	atcattttctg	234660
gaaagggtta	caattttattg	tatctggtgc	attattttctg	gtaccatttc	tggcactct	234720
gtgtaatctg	acattgggtc	gacccagagg	tccaccctcc	gaacttgaca	ttaggccgac	234780
ccagcgggtcc	accctctaaa	ctcgagttag	cagaaaaatt	tttaaaaatt	ttttgagatg	234840
gagatggagt	gaaacttcct	aggtgttaagg	tgcactacag	aggcttgtga	tcctttcttg	234900
gcggctcact	ctctagaaac	gtctgtccca	gaaacacata	aaaaagataa	agtacatttc	234960
tggaccactc	ccattttctg	tgtaatctga	cattgggtcg	accagcgggt	ccaccctcgc	235020
aaactgacat	cgggcctacc	cagcgggtcca	ccctctaaac	tcgagttagc	tgaaaaaatt	235080
tttataaaat	ttttgatgag	gaaaaacact	accagtggag	cgaccgtaca	cactcctcta	235140
gctaactact	aaccgagaga	gtgaggcaga	gtctgtccag	atatctagggt	cgggaagaat	235200
ccagaaacga	ttcaaaccat	ttctggaaca	atcattttctg	gaaagggtta	caattttattg	235260
tatctgggtc	attattttctg	gtaccatttc	tggcacctct	gtgtaatctg	acattggcct	235320
gacccagcgc	tccaccctcc	gaacttgaca	tgcagcctag	ccagcgggtcc	accctctaaa	235380
ctcgagttag	gcagaaaaat	tttttaaaaa	tttttgatga	gggaatatag	tactccgtag	235440
ccaacatata	cacatgaaca	catgaggcgc	tctacaacag	aaagagaact	gatagctgtt	235500
ccagatatct	gggtcggcca	gaacccagaa	acgtttcaac	tcattttctg	acaagccatt	235560
tctggaagag	ggtacaattt	cttataactg	gtatatcatt	tctggtataa	tttctggcac	235620
ctctggtcaa	tctgacattg	ggtcgacca	gcgatccacc	ctccgaactt	gacatcgggc	235680
ctacctagcg	gtccaccctc	taaactcgag	tgacgcagaa	aaatttttta	aaaattttttg	235740
atgaggggaat	atagtactcc	gtagccaaca	tatacacatg	aacacatgag	gcggtctaca	235800
acagaaagag	aactgatagc	tgttccagat	atctgggtcg	gccagaaccc	agaaacgttt	235860
caactcattt	ctggacaagc	cattttctgga	aaggggtaca	atttcttata	actggtatat	235920
catttctgggt	ataattttctg	gcacettcgt	gcaatctgac	attgggtcga	cccagcgatc	235980
caccctccga	acttgacatc	gggcctacct	agcgggtccac	cctctaaact	cgagttagct	236040
gaaaaaattt	ttgaaaaatt	tttgatgaag	gaatatagta	gtatagtgcc	cgccaaagca	236100
tacacacact	cgcccatgct	cgtctagctg	atagctgttc	cagatatctg	ggtcggccag	236160

aaccagaaa	cgtttcaact	catttctgga	caagtcattt	ctggaaaggg	ggtacaattt	236220
cttataactg	gtatattatt	tctggtataa	tttcaggcac	cttcgtgcaa	tctgacattg	236280
ggtcgaccca	gcggtccacc	ctccgaactt	gacatcgggc	ctaccacgag	gtccaccctc	236340
taaaactcgag	tgacgcagaa	aaatttttga	aaaatttttg	atgaagagat	tgagtaaaat	236400
ttcttgacga	taagaggagg	cagtaggtga	ggctgcttgt	ttgatgtgtc	agccacatct	236460
gcgtcataca	ttatatttcc	aagaattttg	ctgacgtcaa	tggaccataa	aaggctttgt	236520
acgtccagag	acaagtttta	gtctgataga	tttcttaaaa	aaaagagggg	gggaggtttg	236580
tttttgtggg	ttctgtgtgt	gtataaaaga	taggtgcaaa	ggtagagaat	catcatatgg	236640
acaaaatctg	tccatgagac	ctcagtagag	aacgcaccat	ggttgcttca	actccgtgtc	236700
caggcccagg	accagttcca	acccaagaac	ttctttctac	aaactttctt	gaagctcaca	236760
agcttgctgt	ggaacttctt	ctcccgtcct	acagtagtga	tgtagtttat	tgtgactctg	236820
agacgtacac	caaacctata	ccgatttttg	ggaacaagag	tatagtttct	accattggag	236880
actatgtctt	atcaaacccc	aatgaagatg	tgagttacca	aatggtttct	tccgtcttag	236940
aaaaatttcc	cttgctattc	cactgcactt	ataagacgaa	tgaagaagat	aaaggatttc	237000
ctctgtggaa	gaagttgtac	aacaaaagaa	aatttcaaact	cctcaactca	ttgttggttc	237060
ataacaacaa	gaactggact	cctgttccag	ctatcccgtt	tgacagggag	aatatatgtg	237120
atgcttcagg	aaggagtgtt	cttatgagtg	aaataatgtc	cacgtcaact	tttcagacaa	237180
tttgcaaaaa	caacacacat	tacttgtttg	atatgtttaa	tatggaacgt	ggcaaacag	237240
gagggagttt	tcttcaactc	tttgcatcta	ggaagaattc	ttttactaac	tttgaatatg	237300
aagaaatgga	ctctcatgtg	ctcagtaaca	tagcgaaatt	catatgcaat	gaaaaggaaa	237360
aactagactc	tttcatacct	gccaacggaa	aaataccatg	ccctgataaa	actaatgtatg	237420
aagggtacat	cccgtctgga	atagcaatta	tggaagacaa	ttaccctgca	ttgctatatc	237480
tcgtttgttag	gtatggagca	tcttgggcaa	acacatacgg	ggatcataat	gaatctctca	237540
aagcgtttgc	aataagaaat	gatgcaaaag	attgtctgga	aattatagag	tttataagtg	237600
atcactacag	tttcaacaaa	aatgtgacga	aggaagaatt	tgtaaagag	aagactgtag	237660
aatgtgttgg	atgtttatat	gatattgaag	acgagaaacg	ttgttacaaa	ctcccatgtg	237720
gacatttcat	gcatacattt	tgcttgtcta	ataagtgttc	taaagctaac	tttgatgtg	237780
ttaaatgttt	ccaaaccttt	gatgacacaa	tttttagaaa	atgtccccc	actatacaat	237840
ggaaaatggg	tataaaccaa	acgactaacc	ataaggaaat	ggatttggtc	aatcgtgcat	237900
ttgacacata	tttagatttt	atttgctcat	ataacgtcaa	attagacaaa	aatcaaaac	237960
ctaaacacaa	acctgaaaac	aaaaaggtgg	aagaagaact	agcaaaaagg	acagcagaaa	238020
ttgaagaggc	tataaagaaa	aaggaagaa	aactagcaaa	aaggacagca	gaaattgaa	238080
aggccataaa	gaaaaaggaa	gaagaactag	caaaaaggac	agcagaaatt	gaagaggcca	238140
tgaagaaaaa	ggaagaagaa	gaactctcaa	aataataata	aataattgaa	aagggaaaaa	238200
gacgactgaa	tgaagaatgt	gtcaagctga	gagatatttc	aactgcagcc	ataaacatgt	238260
acaaagagaa	agtgaagaat	aatgggtgat	tactaaaaga	ttccgatcag	gagttggctg	238320
aggcgaaaga	gaggttgagg	aaaattttat	tgctagaaga	agaaacaaaa	cttgacagat	238380
ttttgttttag	accgaaacga	gtagaagaac	gtatatttct	aactaaagat	gatgaaacgt	238440
tagccttcaa	gttagcccta	gaaaagaaaa	cggaggacat	aattgcaag	aaaaacaacc	238500
aaaaaggcag	tgaagaaga	gatggagaat	atactataac	ttctcatatt	gagaaactac	238560
ctcaatccac	tgctttggct	agtgtgtgtg	tgttaaacga	ataaataaaa	gtataaaatg	238620
taataaatatt	gttttaatat	cattaatat	ttaccacaat	gtctacttgt	tcgaatttgt	238680
tgtcagtatt	tggtggagga	gattggacaa	caacattccc	attcgacctc	gtccatacac	238740
gtcaagagtg	tgataaaaa	agagagcaag	actactcatt	tttcattact	gaaacgtgta	238800
aaggagagaa	tattggtata	cattcgtatg	aacacacgtc	aaagattatt	gacacgggta	238860
ataatgattc	tacctcaata	gaggaactag	aagtactgaa	tatatacaaa	gctataaacc	238920
atttagaaaa	tatcctaata	ctcaacaaag	gagaaaaaat	tatactgatg	gatgtagaaa	238980
caatgatact	ggaaactcat	aaaattttta	tgaagggat	tcttcccaag	ggtaaaaatg	239040
gaagtttcag	tacatgcgta	cgctttgctg	taaataagaa	caatgaacgg	cattactacc	239100
ctgtatttga	aacagagaaa	gaagcgttca	attctataca	aaatctagta	gattattata	239160
atgaaattgt	agctcacacc	aatgacccaa	ttaaaataat	aaaagcgtgc	gcatatttca	239220
tgtacaactt	tctaactctc	caccctttca	atgatggtaa	tggaagaaca	gctagattat	239280
tgtatagttt	tctattgaaa	ggtaatggta	tcgtacctca	tttttcaccc	ataacacacc	239340
ctagggatca	atttgttgat	acttttagtg	atttttagaga	acatggagat	ggacgacctt	239400
tattgtatgt	tttgctggaa	tcaataaaaa	ataagtaaaa	ttcattttga	ggcattttat	239460
tcttcatcaa	ttcatatggc	atcttgccag	ttttctgtat	tgttatctca	gcaataatgc	239520
gtaacacatc	agcagatgtg	cagattttat	ccccctcccc	atcccccatg	gtattttcctc	239580
ctgcaaaagta	atggaggaaa	gaatctccag	taatgggggtt	agttacattc	aatacacgaa	239640
agaattggcc	aatatctttg	tcggccaact	tgtacaaaa	ctccttggtc	atgacttcac	239700
tcataatgat	agattttcca	gtggtgtcac	aaatggggga	tttatcaaac	ggcatggcgg	239760
ggatattggg	gtggaatggg	ttggagaata	acaccaggtc	gccgaggagg	gcaaacttcc	239820

cgtgcttgta aagctccttc cacaggggagc tattcgtctg cttgtagttg caatgaagta 239880  
ggacagggaa tcgcctgagt acgtgggaca gcactttgaa gctcatcttt tcgtccgaag 239940  
aagaagacag tacactctct ccaacatcgg tgacaatatt cttccctccc gacttggtct 240000  
tgtacaatgg agactcgcaa tcggtgaaca tggcgctcgg agggaggagc atggtttgca 240060  
tgaacatgct gccgcggtg aaaaagtcct gctcggcaat ggagaatatg ttttccaatt 240120  
tgcttataag tgagtgcaat gctgctgctt gctgcttgct gctgtaggac gtccagaaag 240180  
cgtgtgttaa tttctcggcg gaggactgtt cttatataca aattcggttaa aggggtgacgc 240240  
ccacagataa gaatgtgggt agacttgaat ctatttgaaa ccatgacaag gcattttctt 240300  
ttgtaagggg tgaataagga ctagactcct ctaccagaat tcatgtgcgt cattgcaagc 240360  
aaaaactccc tcttgttcca agaattggct agtggcggcg tgcattgccag tctgtatgac 240420  
atcatttctt tatgtgcaat ggataattat aatttgtaat atttgaagtc gtcgggtccg 240480  
gtccaccagc aaagggggcg aatacacata acatttctgg ggggtgtattt agtgcattat 240540  
atctggcatc cagaaacgga tgcaagcagt atctgaaact ctctctcact tttgacatca 240600  
atgaaaaatt ggacatcaac acgcatttgg gtggtggata aaagggggca accggctatg 240660  
actaacaagc atttctgccc taacctgacg ccaagcaaca gcattattaca caatggagga 240720  
cctaaaatcc actatcgaga gagtatatga agaaagagtg gagaatctag aacaatggac 240780  
aaatactgta gaggaagaag aaaggactgt ctcagcaatc gattctgtcc tggaggaaca 240840  
aaaaagggcc ctggacgcac gggaagcagc gataaaggaa cgagaaaacg acctcgcatg 240900  
aaaagaaggg atatctgcac tcgttttcaa cgcagcagac gccaaaacac gtaagaattt 240960  
gataaatacg tggatagccg aaagggaaac gtcagaaaaa agaagaaagg aagcaacctc 241020  
taccaataat caactgaaga accagatgtc atctctagtc aacacaacca aaacatcaa 241080  
agaaaagtac aacaaatatt acagaagaag tgccatactc aacatgcaat acatcaataa 241140  
caaaagggat tatgaagcaa gtcaattttg ggtgtatata aacaatgcat aagttttcaa 241200  
ataaattttt ttttataata aaagggtgtt taattataat ttttgtccct gatgttgtgt 241260  
tttctatatt tttactgcct cctcttgggg taagacataa aaacgggtggc ggcggaaacg 241320  
aggaacagaa gagcggaccc agccagaagc atcatctcc tggtcctggt cttatatttg 241380  
tctcatcatc ggcagtgctg tcatcatcat cgggtgctct atcagtgtca 241440  
ggatcgtgtt ccttcttttg tcccatccat acattcatga cggccaggac gaggatacca 241500  
attgagatag caacgatcaa aaagagaccc atgcgagcca tagacatgga gccttcaaga 241560  
tcttccatgg tgtaagaagg gccagcggcc tgggctccgg tcttgccgaa aggaagagt 241620  
ttagtcgtgg tggccatttt ttttctgttt gatgaagagg tacgtcaatg taacgtaaa 241680  
acctgttgtt ttatcgtgta cggtttttgt caggacacct tacgaacctt tttgtctttt 241740  
tagacttact cagcgttagt tgcgcagtt ggtcatgttc cgtcaattct gttcactata 241800  
tcttcttcaa cgccgggtta acgacaatct tagatccacg gcgtccgctt cggcagcggc 241860  
ctcattaaaa ggagatggaa ctgaatttat aaccggagaa ccaccttctc ataaaatgag 241920  
gggaccttct tatagcgttt taggacctga tccgtgcgag gaccagaaa gggatatgt 241980  
tgatattgta gtgtctattt tgcagacaaa taatatcacg gtaacaaaag aatgggaatt 242040  
gttttccgat aagttgagaa aattgggtcc atggattgat aggagcggaa ttgagaataa 242100  
tggcgaagga gaagaagatg gagatgaaaa tgaagacggg ggtggaaatg ggggaagaat 242160  
tgaagacaga gaagcacatc gacgaaaaat gatgaagaaa ttgtcctttg ttggaagaga 242220  
agatccagtc gctgtagatt taccacagtg gcgagaaaac agtacagaat ttgcacgtcg 242280  
tttaacactc aaggaattgt gcgatttaat agttgaatgt ggatgcatca aatcaaaaag 242340  
ggaactcttt gacttcattt ttgaagaacc gtggagagatt aaagaggctg ctgacgttag 242400  
gggtatggca aacaggagta aattcaccaa ggaatcatta attgactggg tttttgagtt 242460  
cgacacatat agtaaatgtg tagtattttt tgaagcagtc aactgggtact tgaaatctca 242520  
agcgtctcca atttcatttg tactagatga tatatattgt tgtgtctttt cctacataag 242580  
acgccaaaac tttttaacta gggcaaaaaa cccatcttta acagtggctt catccttttc 242640  
tcccacgccc gacacaaagc ttttggctat cgcagagtgc gtgcaacact ttttaaaatc 242700  
agacattaat attagccaga tggcattaac tgaaagggac tgcctcttcc ctcttttaac 242760  
tgaaatgccc cgccaacaaa aaaaagtaaa cacttctctg gacacaatga agagacctac 242820  
cttatctctt ctaccttcca cctcctctcc ctcttcttcc aacaacaaga gaaagagaaa 242880  
tactgccgct gccaatattc ttcttccagt gtacaggagt aacttttcta cagcatccaa 242940  
taacaagaga ctgaaaactg atgatgggga aaatgcatca gcctgtattc ttatcgaagg 243000  
gtatgcgaat ggaaaaataa gccctataag gattatggta agaaaaatca ctattattcc 243060  
agaagtgttt aaccatcttt tgttccctgt ctttgcctct aaagacactg gtgcgaatat 243120  
cttatttttt atcaaaatga aatcctttgc aagtgcattt ttactcctcc ctggactttt 243180  
tagacacccc aaacaatttc tcaacggggc gtgcaaatgg atgactctag cagaaaacaa 243240  
catcaacgac aacaacataa actcttccag gatgtggagt tacacgctag cagattattg 243300  
tcctctgggc tattacaccc aagagagccc tcaaccttat cagacatgag gcaattttac 243360  
ttcgactaca aacaagagac tacaaaacgt gcagccatta tactttttaa cactcttttg 243420  
gaatactaca ggacaccttc agaagagtgg gaaattccgt ttaattctct gcttaatgtg 243480

atgaataaca	agtgggagtac	actcattcca	gggtgtcaaaa	taagtgcagg	tatcatatcg	243540
aaactcccat	ggaccatgaa	aacaatgtac	gagattgttt	cttcgcccaa	taataataat	243600
aacaacggag	actactattc	tacatgcagg	cgaatggtaa	tggaatatcc	tatcgggggt	243660
ttattgcaca	cgectgccat	aactaataag	tatccacgct	ccagaatggt	cacctgtaca	243720
aagggcaaaag	accaccagaa	gctatatgac	atctctagac	aaatgtttga	tataatgaa	243780
gcaaatggac	aactctgatt	attattatta	cggttacgaa	aattcccaa	agactaaaa	243840
attcaatact	gcactgatga	ttgggttcat	tttatttatt	ttattaatca	tcggaatttg	243900
gggtttatatt	agatctcgaa	atgcaaccac	ccaagagagc	aaaacttctt	ccccaacaat	243960
ctcctcgacc	ccaactacat	attctggcag	ctcaaccagc	aggggtccag	gttctggatc	244020
tggaaacaaa	cccaaagatg	acacatccgt	tgaaggaata	gacctgggt	tactgtaca	244080
gaaaaaagag	taaaaggcga	cagctcgctt	gccaatgttc	ctgttacgta	ctctgtggtt	244140
tcacgaggtt	gtcatcacca	aaggtaacct	ttttttttgt	cctcgccgac	aaaacgacat	244200
cttaataaacc	aagcaacgtt	cgataaagaa	aaaaactcgt	catggatctt	tcttttactc	244260
tttcgggtcgt	gtcggccatc	ctcgccatca	ctgctgtgat	tgtgttattt	attgtgattt	244320
ttaggtatca	caacactgtg	accaagacca	tcgaaaccca	cacagacaat	atcgagacaa	244380
acatggatga	aaacctccgc	attctctgtga	ctgctgaggt	tggatcaggc	tacttcaaga	244440
tgactgatgt	gtcctttgac	agcgacacct	tgggcaaaat	caagatccgc	aatggaaagt	244500
ctgatgcaca	gatgaaggaa	gaagatgcgg	atcttgtcat	cactcccgtg	gagggccgag	244560
cactcgaagt	gactgtgggg	cagaatctca	cctttgaggg	aacattcaag	gtgtggaaca	244620
acacatcaag	aaagatcaac	atcactggta	tgcagatggt	gccaaaagatt	aacctatcaa	244680
aggcctttgt	cggtagctcc	aacacctcct	ccttcacccc	cgtctctatt	gatgaggtatg	244740
aagttggcac	ctttgtgtgt	gggtaccacct	ttggcgaccc	aattgcagct	accgccggtg	244800
gaaatctttt	cgacatgtac	gtgcacgtca	cctactctgg	cactgagacc	gagtaaataa	244860
atcgtgcttt	tttataataga	tagggaattt	taatattaca	acaataagaa	aataaaaaca	244920
ttgaggaaat	ttataccata	ttttattgac	ctacttaacc	ttcttgctat	acaatgaatg	244980
tttaggtgac	tggaaaagtt	tagcaatatt	atccttgaac	gggaaacatg	caccaattac	245040
aggcgcaatt	tcatacgctc	tcggcctatt	ggctctttcc	tggtcataca	ttttagatac	245100
aatagacaaa	aatggaatgt	ttgtatagat	agaattggca	gacaaatctg	cagttctctt	245160
aatcaaaatg	gacaacatgt	ctattaacaa	ataagccaac	ccaaaagtca	tggcagtttc	245220
tgaacacaaac	tcactgttaa	taaaattcagg	agctgtatga	ggatggttac	taaagaacct	245280
ctcatcagtt	ccccaacatt	taaaattgta	gtacttttta	catggtacaa	ttaaaccaaa	245340
atcaatcatic	ttaggttgac	cagttattcc	atcaattact	atattgtcac	tttttatgtc	245400
cggtattcact	aatccttggt	gagacaaccg	agtgacaata	tttacaattt	ctaccaaaaac	245460
aaagggcaat	tggtttaaca	ttctctccct	catttttcca	acgatagcta	tggctgaaat	245520
tgaatccgta	atgggtttct	tgcattttaga	ttgtagacct	tcaggcaggc	ggccagtagc	245580
ttgaagcatic	ctaacaccgt	acagagtatc	cctcatccta	cacaagctgc	catgatcttg	245640
catgagctca	gcggctggct	tggaggaagt	agtcgagggg	ttaggttgct	gtacaacatt	245700
ctcgtccatt	ttacatttca	gggtctccatt	taatacatca	atcacagaaa	tgcttgcaaa	245760
atccatctcg	acacaaaaac	cttcaacaca	catctttctc	acacctactg	cgcctttcat	245820
tcttccctca	agagcgtgac	taaaaacagt	ttcaaataca	aattcacagt	acatttcctt	245880
gtccgtcata	aacttgataa	cttttccagt	ttcaagcatg	taatatatcc	cgactttaac	245940
ttcgagatat	aattcagatg	gtccaatcat	tacgacttga	gaagcatatt	ctggttggaac	246000
gtgtgctggt	acaggaatat	ttctgcggtg	actttcccat	ttcttacatt	gggcagtgac	246060
ctcccccaag	agtgtagttt	tcttctccaa	tatagattca	catgctccag	gggaaagctt	246120
aaaaacattc	tttgccatgt	gatcgatgct	catggttttt	atccaccctt	taccattatt	246180
cttgacataac	ctgttccagt	tcaagagagg	gctagtgtct	aaaacggata	ggtttttact	246240
accaaaccca	acaggccacc	aagatgaagg	gggttggtta	gtacggttgt	tagaccacaa	246300
aaacttggac	gctttttcct	ttgcgtcaat	acagctctgt	tttatggtct	cgtagttgga	246360
cgacacatct	tgcacgtcag	aaatgaggcc	taattttgct	gcctgtctaa	tcaaatggaa	246420
cgagagattct	agaccgctcg	caatattggt	atcagcctct	gccactgcaa	agtaacacgg	246480
gggtttttct	agcttgaccc	acttgtcaat	ggaagaaaag	ttcctcttta	ctacatttgt	246540
ttcagtcacg	gtatcacacg	aagatttcaa	gtgttttggc	cccttgacat	gggaacggag	246600
gacagacagt	ggagtatagt	ctctagcctt	caaattgaca	gtagagaatt	ttgcacaatt	246660
tggatgagag	aatctttttc	tgtactgggt	gtcggcattg	gcgcggtgct	tcctcaacac	246720
agacgcggga	ctaacagaag	gcaattgctg	cttcttgata	atgtctgtcg	catacgttgg	246780
atthaatggt	gcaaattgat	tttcatcaat	agtaataggc	ctcttcttct	tatttctatt	246840
aaagtgttaa	ggtgtgctct	cgtagtctct	gcggagcctc	aaagtcacag	gaggcggtgc	246900
ttttctctcc	acctctccat	tattcttctt	tagtctcttc	atgggcctaa	gagaataagg	246960
agaatctgag	ggaggggagc	ggcgagtagc	gttttttagag	caacaagatt	tcacagctat	247020
ggctgaagggt	atcttgaatt	gccctcctcc	ttctcctcct	tctccttctc	ctggcggtttt	247080
cgattggtaa	agtcccatca	cgggttgctgg	cgtaagtttt	gtccgttggt	ccccaccctc	247140



cattttaacg	acgtttttta	tactctgctc	gggctctggg	tcagaggggg	ctatagaaat	247200
cactcgtgga	tagttccttt	atatatttaa	ccaatttttc	aaacaggggt	tattaaaact	247260
actcttcttt	tctgtatata	ttaggtacaa	acacacaata	caggtcaggc	ctgcaaagga	247320
ctgggtattt	accagtcoga	atataacagt	caaaaggctt	attttccttg	atctagagct	247380
ggaaatgagt	gagtaaaaac	ttgctgagac	atatactgtc	tggccgcccc	ttccttatta	247440
gcctgaatag	aatcctcgaa	gcttacaact	acattaaatt	tccccatacc	gttctggggt	247500
ttcaacatat	ttgcacatcc	aaacgttaca	gttttcatgg	atactacacc	ctcgtctcct	247560
aacaaatcac	agttcctact	atttttgtca	gccgcattcc	tttgccctct	gacagattct	247620
gttacagcgt	cctggatttt	cttggtccac	agagctttaa	tagatctctc	ggacaattgt	247680
tcaccatacc	ctaaaacagc	acctctggca	gtaatggggc	agtaattaaa	acagacaaca	247740
gaacctcctt	ctctccaaat	tctggtagat	agattgtccg	cttctttatt	caacaataaa	247800
cttcttagac	agtttgaaat	aatgtccaat	tctttactct	tgtgggaaaa	gggtcctcca	247860
ataaatcgac	caaacacacc	cgcaccagaa	gaaagggtct	cgtagaaga	tggttcatta	247920
aacatcatat	aaaatttgac	atcagaagga	ggtttcccg	tcttcaaaaa	atcccagaag	247980
ttagcagcga	tgggtaaaat	ctgtacaggg	ttcacctttg	aagctctaac	ttctattaca	248040
gcatttgata	tgtcagtttc	catggccttc	atataatttg	ggatattcct	attaacaaaa	248100
tcacacatag	gtattttcca	ttcagggttg	atgtgggttg	aaagaacaag	ctctagaagc	248160
actttaaaga	tgtaatcgcc	agtgattgtt	gtgactggag	tgttgattgg	gtatagagat	248220
tcttggtgct	tcttttccct	ttctgatgct	tcatacaaaa	catccatgcc	ttcacagtca	248280
tcgtcatcat	catcatcatc	atcttctact	tccgacattt	ggtagggcact	attattatta	248340
gtgttcttga	cactgtttacc	aattttcttg	tacatttttt	caatagttaa	aacattatga	248400
ttatcagcac	tctcattctt	cattatgttc	ttcagtga	gatccctctc	cacgttgttg	248460
tgtttggaca	cgtaagcctt	gacgatcttc	ctccatgtcc	ttataaagtg	tatgactgat	248520
gcataattct	tacccgtaac	aggtttatga	gtcagcgagg	cttggttagag	cttgaaggcg	248580
tttggtccagt	gattatatac	agtcgtgtac	aaggcctcga	tggtttcctt	tgccagctca	248640
gggatgtag	ttttactccc	atatttcgaa	tccatagtga	gtagtctctg	tttgttgtcg	248700
agaaatgtcct	tcaataaagg	gtccaacgta	ctcttattta	aaattttgtg	ttcgaattca	248760
tatatcagcg	aacaagtttt	catgtctttg	tatgacacac	aacatctcac	aagacagatg	248820
actgcagtc	tcagaaacgt	atcacttttc	ccacttcttt	ctttcttcgt	gtcaagaaga	248880
ggataaacgt	tcttgtacca	gtaatcgagt	atagattctg	cattggctgc	gatggtttct	248940
ttaacagctt	gggttaaatt	agatacagtt	ctgtcctctc	cttcaaaaat	cactttgtct	249000
tcttctacct	ggcggcattg	tcttttcttt	gtgtttttgt	ttccagagag	agtaattgca	249060
gaatttttct	gttcaactac	cagaactcta	taaagatttt	ccactttttg	gaagagagat	249120
ttctggcggt	catcttcaac	aactggagcc	atgtctgtac	cggagccaac	tttgttcaca	249180
actaacgcgt	cgggtgaatgc	gtctttttaa	gccattacgg	tccatgccat	atctcaaagg	249240
ttgccaatga	ccaacataaa	tatcggcaca	gcatacattg	aattgcctaa	tactataaag	249300
gaaggactcg	ataacccgaa	acctatatac	tttgatcctt	tatctaataa	atataaacgg	249360
gatacttgca	ctgacgtgct	ggactcgttg	gacgtgtcgc	tagaggacta	tagaaaaagt	249420
ataggatggc	gtcctcgggg	ggattcttta	caggaataga	tgaccttttc	aagacagtga	249480
ttcaacaaga	aaaacaagag	aaaaataaac	caactcaagc	accagaaaca	gaaccaaacc	249540
cagggccatc	tcaagctcca	gatccagttc	cagaccagtt	tcctaaaaca	ccaaccaatt	249600
tctgtctctc	tccacctaat	cctctcctcc	ctctcctccc	tcctcctcct	aagccctcaa	249660
gagaagaacg	gctaaagacg	tcaaaaaatac	gtttaaacaa	agctcttagt	gatattgttg	249720
aagccacaaa	cgagcgtggt	gatgcgttga	aagagaacca	agcattaaat	acagaatatg	249780
acaagaagga	taattacttc	caggttttta	agtgtctgat	aacaccttct	gtaccaacag	249840
ctatttatagg	cgcacacgtg	aaacaggttg	ccaaaagtag	cgaaatcgaa	ctggccgtga	249900
acgaactcga	tataaaaaat	aagtgtctct	tagtgtacaa	cgaaaatgag	tcgttaaaat	249960
ttttcaggga	ccatgagaac	cttatactac	aaattgccgt	ccagttattc	tctaggcacg	250020
ataacaccaa	atgcgtgggg	gcagaaatat	gtgttaaagg	caacgaaaaa	aacaagtttg	250080
ttaacaaact	ggtggtaaaa	aaactcccca	atgcaccatc	ctcatcttca	actgtgctag	250140
aaatttagagg	cgctaccaga	aatttactgg	agaataattt	caacaaggga	gaaaaataca	250200
ctgtcaatga	aaacaaggag	attcctcctt	cagaacgagc	caacctggac	acgaccaagg	250260
cagaaatatac	gcactgtctt	tccactctac	acagactgga	cactaaaagg	aagcttttct	250320
ttaaaggcaa	cactttttat	caacgaaaaac	caacattcga	taataaattc	aggtggacag	250380
aagtatatagg	gtggacagaa	agtgaagcat	caaaacaaac	cactaaatcg	ctagacaagc	250440
caacggacga	caatttatcc	gtgtaccccc	attctttcaa	taatttggca	gaccacttac	250500
gtttgaaatt	taaaaacgct	ctctataaaa	atagtaccgc	acatcccggc	aaacgaaatt	250560
actacaagac	tcaagagacg	ctaataaatc	cccagattga	ttcggcgaaa	gagtacaaga	250620
tgggtctttgc	agaaatcgac	aagtgttttg	atgttctttt	ggccataggg	aagaatgaca	250680
aatacacaaa	aagcactgtc	atacaatata	gaggaaagtt	tagaaggtat	ttaatatctt	250740
gctacgcctt	ttatgctcta	aataaggcaa	aacattctcg	cgcagtatcc	cctctaccat	250800

ttaattttctt	taaccttttc	tccttcatgt	attgtcatgg	tccgtttctc	cattccgcca	250860
gttttttgtc	cacattgacg	ttcgtctatc	aacacatgtt	tttccccatg	ggcacagccg	250920
ccccatccgt	ctcagccaag	cggctcatgg	atatacgattc	cgccctaata	aaaggaggaa	250980
aggggggtggg	tgtgagggat	tttggttcac	cttcaaaaaac	aagtctccat	acaagaacat	251040
tgggtgtcttt	cttaggtttt	gctgaaatgg	ctatgggaac	aatgacggct	ctcttatctg	251100
gtgtagaagt	gcgtgtatct	ccagctctcc	aacaaaggat	atctaaatcc	ctagaaagat	251160
ggtgtgatcc	agtcactctt	atatatttca	cctttgtttt	attccacaga	ttcagtgggtg	251220
cgaaaaaagt	atcactcgaa	tcggcgcttc	gcctcatcat	ggggcagacg	cacgcccaca	251280
caaataaggt	gagggccgcc	aagagatgcc	gaatagaagc	agcggaaatg	gaagggtgtg	251340
aagaagaaga	ggcgggcctg	acactctctt	atgccatctt	attgggtctt	ccttactcta	251400
tacaaaaagc	cctcggatta	cctgtcccta	agataaaacc	tctcatgaca	gcactcttct	251460
ctcaatacaa	tttaggggat	ttttagggcg	tggaaacaact	tctaaaggct	aagagagagt	251520
ttccagccga	aggagaaacc	gcaggatttc	tcggcatgtt	tgataatcta	gtgaaagatt	251580
ctattgacaa	atactacggc	gaaggagcct	tttcagacgt	ggttgaaaaa	gtaaaacaag	251640
gcattggaaca	aaacacaccg	tatgacacat	cttcagcgtt	gatgacacct	atccctaaag	251700
cattctacga	agaagaaaag	gatgttccac	agcaggaaga	aaattctaca	caacaaagat	251760
atagattgaa	tagagacgtg	gaggaatatt	taatggcttc	tcctatgaag	atggtgtttg	251820
tgtctatact	cgataaaaact	aaccaaaaaag	aacgtttcat	gtctgttggg	gatattgccc	251880
ttctggccgt	gtggtgcaaa	aggaacgtac	tgaaaaaagga	ttggaacgaa	tacgctatcg	251940
ctaaaggcaa	ctacgaatgg	cttggtgcta	aaatgtgcaa	ccatttactt	ttagctgatt	252000
tagtgaattt	tggaaatta	ggtgacttga	aaataaccaa	taaacttgac	acaaataaccg	252060
acacctttca	cagagacagt	gatagattac	cctcagttgc	agatcagaaa	aaatttataa	252120
aaaacacatc	cctatctgat	cgaaaacaat	tggcccttgt	tcactcgtgc	gttaacgtga	252180
gcacccgaac	ccacgtagga	agagtgactg	caacatcatg	ggctgtcgat	gcgcttcgta	252240
cctatacaag	aggtgataaa	gacatgtttg	ccgccttatc	ttcatcgctg	gatattgtacc	252300
atcttgggca	cacgaattca	gctaattttg	ttccataatt	tagtagaaat	tacctaagta	252360
acgaacaaga	gaatggattg	tgggggtata	ctcgcagaa	ctctgaaaaa	ttggccaaag	252420
aagaattggg	aagaggacgt	ttagggggcc	tgaataaggt	aggggtggct	aaaacagaac	252480
tggctgctgc	agccattgca	atttcttctg	ccttagatat	gggggaagta	gaagctgtaa	252540
tggacgactc	ttctaaaagt	agaaaaatag	cctccacctg	cttaaatgtt	aatgcagcca	252600
aggtctcggc	cgccagagaa	aaggcgagag	aagctagtat	taaacgtctt	cttctggcca	252660
ctaactgcacc	agcagctggg	tcattccagaa	acagttaacag	gtttctcctc	aaagatttgt	252720
gggggttctt	ttctgaccca	gacaagcgcc	agaagcttat	aaagggtgaa	gcagtttctg	252780
tactatgtcc	caatacagga	tttcttcatg	ctgctgttcc	tgatttttgt	attgagtatt	252840
ccttcgaaaag	tgaacctctt	atagttagat	tacgtttgag	actgattaaa	cctgaaaaac	252900
aagacgaaat	ggtatgccct	tcaacagctc	ccgaagctaa	taagaagagg	aaattagtaa	252960
ggaataatca	agacgtgtga	ctgacgttgg	atgatgaaga	taacatcggt	aaatacaaca	253020
aatatgatat	ggttgaagac	gaggaagcgc	gtgaaagatt	acgccaccag	gacaaacaat	253080
cggttatttg	agcccgatc	agtaaagtgt	gtgagcggaa	aaatccaaag	aaaaaacgtc	253140
gtttagaaga	ccctgaattg	caaagtgtgg	atgaacaatt	gatacgggaa	ctggctgcca	253200
ttgcctactg	acgagatcta	agattttgta	atataattatg	tgtcttgcag	acctaataatc	253260
ctgaggattga	ataataataa	caacaactgt	tcaaaaatga	gagacgatac	ttttaaccaa	253320
gaaactgcgg	tgaacttgt	acgatgggat	acagagtacg	attgtttgtt	cccattgggt	253380
aaccgcgtgg	agcgccttct	aggatcggtc	ggaggaggcg	tggacgccac	gtctgtgcga	253440
agccgaccgg	ctctttatga	agaagataag	aaggagagata	aatgcatacc	ctttaggata	253500
acgtccctta	ttgaggggat	acttttgtaa	agggctctaa	ctaaaccgga	tttagctgct	253560
gcagcttttg	atgtatcaga	aaagctgggt	tattgttagtt	gtaataacac	tcaaggcaat	253620
tttgatgtct	cttcaatgac	catatggatt	gatggcaata	atagtaaaaa	gtatgaagtt	253680
acatgcccggt	catgcactgt	cgagaaaatt	agtggagggtg	ccgaatctat	tcacaagaaa	253740
cccatgtctc	ttcttgccct	ctttaacaat	ctggtagaga	aagaagcctt	cgccgaaaga	253800
attgaactca	agaaattgta	cctctcctta	ctaacgggct	cggcagccgg	aggaggagggt	253860
atgtacaagg	acagctccca	acaattcttc	ttcaacggct	cttggacgtc	gttgttgttc	253920
cacacatcta	aaaaggacaa	gactcgttta	gaggctgaag	ttttagtcag	taacaagata	253980
aaacacacat	caagattgca	gcctaggtgc	gtctgttccg	atctgtttata	tgccctatgc	254040
tccaccacta	acaactctgc	atcttacgcg	tacaaggcaa	gaaatttgtg	tgttattgaa	254100
ggtggggaat	ttttatat	ttaatacaca	atctttgaag	agaatggacc	tttcgactcc	254160
aaaaacagacc	ttcaatcatt	agttaataac	gagcctgttt	ctgagacaaa	ttcatcgcca	254220
ctggccgctt	cttcttctc	tttagaagac	gacgatgatt	gttgtgatga	tgatgacgat	254280
gatgatgatg	atgaagacga	aaaaactaag	aagaaacaac	ccaagaaaca	aaccaagaaa	254340
caaaaaacaa	caacatcaac	acttccacct	atcagcaaaa	ccaatcacga	caacatgttg	254400
atgaatgtac	ttaaaaaaag	agctgttaat	ggaaaaacgga	aaatgatgga	ttctttgtca	254460

ggaaaaaagg	gccaacactc	taaaaaattg	aaaacctccg	ctgctgctgg	tgggtggtgct	254520
tcatccgacg	ttgttgacag	agaaaatgag	gaagagaaca	acccctcttc	agtgagtcct	254580
actaacaata	gggatagaaa	agactatgtg	cttccatgcc	ctcaaataga	agaagtcact	254640
atTTTTTcac	aacacaggat	gaacaataac	aagttggcag	aaagtgtagt	caaacattct	254700
gttggttatta	atggaaattg	tttaaacttg	ttgtttactc	aacacagaaa	aaagtatatc	254760
ctgcctcacg	aaaatattct	TTTTTgcccc	ccttttagtcc	agcatgtagg	atTTTaaacaa	254820
tttcgcattt	tgactggcgt	ttcttgcttt	tttgatagaa	ttgaaattgt	TTTTTccgac	254880
caatctgact	ctgtggtggt	gagtaataat	gctgcccatt	cggctattct	aaggctattg	254940
tcatacataa	gagaaaactc	attgaagcga	agtgtgagga	ctgcttcggt	gaaaggaaatc	255000
gattttgtcg	tgaaatcaca	ggacactaat	ataggcatac	ctTTTaagcaa	caaggaaata	255060
agagaacggc	aattatgctc	agcttcaacc	ctgagtatgc	tagctgggtt	gggaaaatga	255120
tcaccgaccc	gggggtcatt	ttaccggttt	ccaaggatgt	tgtactcttt	ggatcacgag	255180
gacaatcaga	tgtgggaata	atgacactgg	accctcacga	tttgatata	aaaatcacat	255240
ctaaacgcat	aggtgtggaa	gaaagactag	ctcaatacaa	tactctacct	atggatttta	255300
cacgggcaat	ggaaaaggaa	ctaaataata	gtagaaatat	gaaagagtca	atattcacgg	255360
gaatatTTTT	agacaccggt	tcggcaatct	tcgaagacaa	catgttcaac	ggaggaggtt	255420
cagctttgcg	cttaattaga	ttccccgctt	tgaattctgc	tgtattttca	agcaagaact	255480
acatcatcaa	acaattgcc	accataacca	aatctctaag	gagaagtcaa	gctagagata	255540
agcaagtgga	taaaacaaga	gagaagatag	tggtyggattc	tttcagcata	cttagtgcta	255600
tagctgctca	agtaatgcac	ctcacagacg	gagagatgac	gtacgtcccc	gatgggcact	255660
cggttaatgt	tgtcatgtca	gagaccaatg	cttcgtccat	ctacttgatc	ataaacgacc	255720
ccactggttc	gggatggaaa	attatgccta	acaatttcaa	taaaacactt	gaaatgagag	255780
acgggtgtaat	agatagagta	gaaacattag	tggagtgtgc	gtgcaagtgc	gtcgcacatc	255840
ccttgattaa	aaggggcatg	gatttagtg	atatgcaaag	aactataagg	tctatggatt	255900
tcctccctcc	agcttcttct	acttccaata	atactcctag	agtagcgata	atgacatctg	255960
gaagttagtac	tactacgggc	attggatcct	tgtccattct	tgcagaagat	ggatcaacac	256020
accaccaa	caagctatcg	gaatatagga	ctgtattatc	cattactgaa	aataatagag	256080
aggtgtcttt	tacggtagaa	ccttcaatag	acggcggttca	agcagagcat	cctctatccc	256140
cttctattct	tcagtggtta	cctcctctag	ttaaaaggcc	agaagtggta	gcagcagcag	256200
cagcagcagt	agtagaagaa	gaaaacgggg	acaataaacc	ttctgataaa	gataacgaag	256260
acaagtacag	tgatactgat	ttttggtcta	atgtccccgt	cacacctcta	attacaccca	256320
agaaatggag	agcgtgcaaa	ataaacgata	gggcaatgat	tagtagttgg	aaaaataatc	256380
tagtgaaact	ccacaaatat	gattggacaa	ataaaactac	aaaggttgat	tatttttgata	256440
agatggctgc	ctttgttgcc	ctcatgacct	ttagaaaatt	ccaagacata	ctagcggata	256500
actatgttcc	tcctcaaacc	ccttctcagg	gaagtgaata	cgcagtgacc	atgtctaacc	256560
tggctacact	ctttactgac	gtgtacgggt	ttgaaatcgaa	tggaaaataag	ccattgtttg	256620
ccctagaaca	gctagaaaat	gaaaccggga	ttgaacgcat	atactccta	aataatcatg	256680
gaaattcccc	tgtatggtaat	tctgtcagg	tcgtcagact	ggaaaaggaa	atgagtttcc	256740
tcttgaaggc	gaagcagtac	tttacagaaa	tggccatacc	tcctattaat	gaaaaatgca	256800
aatggacaga	taaggccccg	tcactctgta	aggagtacaa	gtatttttgt	gatctaaccg	256860
caccatttcc	aaagagacct	agaaaagata	acaacgacgg	cgggtgtggag	cattctgcgt	256920
tgacttatata	acctaggtgc	atataccaca	ctgaacgttg	tttagtccat	ctttactctg	256980
agccagaaaa	aataacagaa	cacgtatctt	tcaacaagga	tttgaacata	ttagaaattg	257040
gaaaaaata	taccaacca	taccaaacaa	actacaaaag	catattcgaa	attgtggacg	257100
ttcccataat	tgtcgctctc	atgtcatcaa	caaaaacaat	gactgtaaac	aactacataa	257160
tttcaacacc	ttctgccacg	accaagtttg	ttcaggatcc	gcaaaaaaca	gggaaacaac	257220
ttctggcagt	tgaagaggtc	agaaacttta	aactcaaatc	tgttcttggt	cctcctcctt	257280
atTTTtaggga	caataagcgc	aacacaactc	tttgttcaca	aataactgaa	caaaattgcc	257340
cgctcatcttc	tgaaggtggg	cgtttttcat	gtccatcaga	gtcacttatt	ctcaagtact	257400
ctaactctctc	taaaaagcgc	gcactggaag	aaattgcccc	agagactgag	actagcattt	257460
tgtcactagc	catgtaacgt	gtctcgatat	atgacaattt	aaacaacctc	ttataaaga	257520
ggataattat	cataatagaa	aattaggtaa	aatgggagta	cagaagaata	tcctggtcgg	257580
tggtggtggt	ggtgtatcat	tgttattggg	tgttgttaca	cttctaggaa	cagtaacaga	257640
aggagcaccg	gcagtcccac	ctttttctatc	gtcttccat	tcctttactc	ctgaatcgag	257700
tgttttctgg	gtcgaaggaa	atcgtgtttt	aagtggaaac	agaaggaca	ccttaatcaa	257760
cgttctgggg	aagaagattc	cttattatgc	taattccata	ttcagacatg	actgttctga	257820
aactcgttct	attcaatggc	cagaaacttc	ccctttgggc	ttgaacctta	ttttctgttc	257880
atgtgcgagt	catgaacatc	aacaccgtac	tcatgaaaca	acagaacctg	acgatttatt	257940
gtgggacgga	tcaagaaaaa	ctaccacat	aattctacct	aaaaaatggg	ggtctgatgt	258000
tgtgtggaca	tctttatgga	gggataacga	ccagaagtgt	gggtgcgggc	aggcatttgt	258060
ttcttctctt	acttctacgc	agaaagaagt	ccaaggggaa	tggcttgctg	ctcacactaa	258120

cggaagaca	tctgagggtg	atactaattc	agcctacett	ttcataagcc	ttcaacgaac	258180
tacactcaag	cctatcatca	ctgacgtcac	agaagataat	atgatgatgg	gaagaatgtc	258240
aggcacaccc	atgaacccta	aggatatgac	atattttgtc	aacgattttt	cagacgatat	258300
aggaagtact	cctcagtgtc	ttgtatcgaa	ttcggacatc	ctgaacaaga	gggaagaatg	258360
gatactgttt	tgggtgtgtg	cagactctaa	agacctccta	actaaacatc	aactggggga	258420
gcgggaatat	gggagcgaag	gaagaagaag	aaatcccggt	gttgaggagg	aagaagaaga	258480
gagagtggaa	gaagaagaag	aagtagaagt	tgcgctacct	tacattaaga	aaagtggaaa	258540
acttatcgga	cctcgtagaa	gacctttgac	aacaacaaca	accactacta	ctactactac	258600
tactaatcct	attgttagag	aggttgtgga	agattttgat	tacgagtctt	ttaatgaacc	258660
agaaatcttt	ggcagtaact	caaaacttcc	cttcattaga	tttcttgatc	aaaaaaattg	258720
gagacttggg	atcatgagca	gagtttcttc	ttccatcgcc	aactttaaaa	ttgaacaaga	258780
gtcatcaaaa	gccttatttt	gtttggcagt	ctgggttggg	gatgaacata	cccctaaatt	258840
cagacttagt	gtatggaaga	actggaagcc	ttttacttct	gcacctatta	ttgtccagaa	258900
tgtaggttat	tcctctgatg	ttttctggca	tgaactctt	agaagcaaaa	ttgttgatcg	258960
gtcaagggac	ctgatagaaa	caaaagtgc	aaagaaaatt	ggggaagatt	gggctaataa	259020
aaaacaaact	gtagtgtgta	tgtttatttc	aggtattgtt	tgtataacag	taacagttat	259080
ttctatatatt	tcaattgtaa	tatatataaa	aataaaaatg	cctaaaattct	aatacaatgg	259140
tttattcatt	ccttattcct	actactagt	tatagcacat	gaaacaaaga	aagagacacg	259200
cttcccttgt	tccttaacaa	atatgtttga	tacatttttc	acaataaatt	tgtcaacttt	259260
atccattact	tcattacctt	ctcctgatat	cttttccaaa	tccaccacag	ttccaacacc	259320
agtaaacatg	ctctttctga	tatccaatgt	tgaatacttg	gcaaattcat	gcccttgcat	259380
ttcagtcagt	gcaatgatta	tagcaccata	cattttacat	tttttgaata	gaggatatgt	259440
ggcaaagtgt	ggagagaata	tctttgtggg	aagaccttct	tcttctccta	aagtactttt	259500
cagtgtctct	ccagttctct	ttacatgcaa	agaaaattgt	acccttttca	ttggatgaac	259560
ttgtccacag	aacaggacat	cttcagacat	tgaagagtgg	agattgttag	taggtgaaat	259620
tgtctttctt	ggtgccaaaa	gactaatcgt	atcagatggg	gtttctgggt	tactttcaag	259680
aatacttttg	tccttaattg	ggaattcact	aacaccagaa	gctgctgcta	tttttgacat	259740
gaacattaac	atacaatata	cattagcttc	gtctgacata	cacagccatc	ccttttctgc	259800
atctacatgg	atattttctc	ccttctcttg	agctgtgcat	gtggacgcta	caatagaccc	259860
atgtttcgat	agaccacaag	tttccgtagg	acgacaaaat	ttaccagtc	tacgcaagtt	259920
gtgctgcata	ctcttacata	caacatccat	gtatttatta	accttatttt	gattagattt	259980
gtcacttaact	tttttcaaaa	taccttcaat	gttatttgaa	aactcatccg	cgcttggtat	260040
aggaagacgc	aatttgacat	ctctttctcc	ccgctccttg	caatgggcaa	aatcttttaga	260100
ctcgtaaccc	agatttggtg	catgagtaga	attgggtggt	attgattttg	ggcgtacata	260160
ctgattcagg	cacggattgg	gagatggcaa	aaaggcgaat	gaagaattca	aaacgttcaa	260220
tttttcatac	actaaaaata	tataatggag	ataaaattgg	aacacaattg	tagccttggt	260280
atgggggtga	atatcaaaaca	acaacatttc	aatagtggac	tcctttgggtg	ccctattcac	260340
aaaaactggg	tgtattttga	tatagtcttt	cttgttcttt	tgtttcgtct	catattcttc	260400
acacaacttt	gaagagggtc	gttttacaat	accactggct	atagagtagt	gtgtgattga	260460
acgaacgata	gcagtgtata	gctcaatcat	tttcttggtg	cataatttct	cataatgttt	260520
agccacatct	ccattattaa	caccaaaaga	aaagtctcta	atgtttttta	gttctgtttc	260580
tagcccatgt	tgtgaataaa	ccttagcagt	gttggccatc	caagacgcat	taaacgataa	260640
agattttccc	gatctactta	ggatgaattc	ttcagcggca	gctaccacat	ctttagaagg	260700
gtcataggca	gaatacattg	ttactcctcc	atctccatta	tagtcatcat	ccattgcagc	260760
ggaagattct	gccacctctc	tagagacact	aactgtatcg	tggttataaa	agtcgctccc	260820
attaaaatca	gcaatgcctc	caataggcat	agtataagcc	ttgtctttac	tcatggcgca	260880
tttaattgta	cacaagttac	cccgttcaac	agaatgattt	tgaataatga	acaagttttg	260940
aggattacta	ttgcgggtta	cagcaacata	cgcatcctgg	gtagatgcac	cagaaaggtc	261000
aacaatagt	tctctaaaga	aagtgtgccc	ttggcaagaa	tatatcgtca	tggcttgact	261060
agattcaaat	ggaagatatt	gaacgtatac	attgttccca	aatcctccat	tcttgatatt	261120
tggattaccg	ataatttgct	gtccttcttt	aatttctagg	cattttttgtc	ctagccgttc	261180
cacgtacaca	aatactgtca	attcaccatt	cttgtacatt	aaattagtaa	ccactccagt	261240
gtccttagtc	acgaaccttt	cctgagtgcc	atggatcatt	ctattagatg	tagtgaataa	261300
cacatttttg	ccctgataaa	gagtaataga	tttagcccta	gttaccttct	ttttctcctc	261360
caaaagaatc	gtagttaatg	ccctaggaga	ttcggctgcc	ctcaaaattt	tcctgtctgt	261420
aaacttgta	aaaagtggtg	tactgcgagt	tttaatggcg	tttaattaat	gttgccggta	261480
cttgtgagag	atggttttct	cgtggctaatt	gtcagacaga	tccatattca	aaactacaaa	261540
ttttacttcc	actcttaagt	atlttcaact	atlttcaact	gccgcagccg	ctgcaagtga	261600
agtagaagag	gaggatgagg	aattgttatt	gttcttttta	gtttctttga	tgatagaagt	261660
atccactaaa	ttcgagactg	cattctttta	aacgtgtttt	ttagttagtg	tgattgcacc	261720
catcatcttg	tcagttttgg	aagaaattgc	atccttttct	tgtttggtct	tagagttggc	261780

catgaggcgt	tcacacatat	ccttaggctg	cccttggtct	tttggtataa	actcttcaat	261840
cttcttcctc	ctctcatttc	cctttgaaat	actagtcacc	atttgcatgg	ccaacactat	261900
agatttttca	gagtatttaa	gtgagttatt	gtgcctattg	ttaatgttgt	tatttttcagt	261960
gtgagcagtt	tgaagcactc	tggaaatctt	taaactaccc	ttaacaccct	ccatcatttg	262020
accattattg	aaaacatctt	cctttctgta	cataagaaat	tcgttaacca	ttgaaaggga	262080
gatgttatca	aaatattgca	ctgggtcaaa	gttggtccatg	aaaacactcc	catttctgga	262140
tgatctaccg	gccattcgcg	tatctaccaa	gtcttgaagg	ttggcagtc	tatcaaaaca	262200
tgaaataccg	tacattaacg	aataagcgtg	cttaataatg	tctattctct	ttaacgcagc	262260
catgatatat	gttctcgaaa	gggcagaaa	ataggaagaa	tatgcaggta	ataacatact	262320
cttccctaca	tctcctcctg	ttgatatggg	catagataat	accctcaacc	aatattcgct	262380
atggagaggc	ttatctgccc	ctgctcctcc	tctcctcctc	actcctactc	ctgtgcttgt	262440
ggcccattct	gagggcgagt	ggaacaagcg	tccagacagt	gcagatttat	caagagaagg	262500
gagttttcca	ttatcaaatt	taacctgttg	cattctcact	agtgtataca	aataccacac	262560
acgggctcgt	accaaactctg	aaggcttatt	ttcagtgagc	cattttttcc	acacaacatg	262620
aaagtcaact	aaatgacct	cacgatctag	agtagatgcg	tcacgggtgca	atttatctgg	262680
tatatcagtt	ctatatatca	tttctccact	ttcttcatca	tcgtcatcac	ctccccacc	262740
actacttttc	ataccagaac	ctccacttcc	tccgtcttcc	aaatccattg	cttcttcttc	262800
atcgaaataa	caatagtcc	cgctgcctc	ataataagct	gctgctgcag	ctgtacgtga	262860
tgatgcacta	ttattattac	cagtttcttt	cttcaagaga	atagatctta	tattttcaga	262920
tttcgctagt	agttctttgt	atgttaaagt	ttcattctct	aaaaatgaag	caactgtacg	262980
caaattatct	ttaaagtaat	tagcgatagg	agtttcatcc	acaaagtctt	ttccatctct	263040
aaagaaggcg	ttctccacac	acgctgccac	agcagacgca	cggttatgtg	cagaattgaa	263100
tggtactgca	tctatacact	tgggatcaga	accgtgcaca	tacttgacgt	tcaatatctc	263160
ttcatccatt	aaggagggtca	gatgtgtatc	cttaaatctc	tctgataggc	tagaaacaat	263220
atacaatttt	tctcgccac	tgtaatcaat	ttctctcatg	gctttggcca	tcatgtcgaa	263280
cacaatactg	catacatttt	cagtgacttt	tcttagattt	gcattgctca	ttttcttgag	263340
ttcagatatg	tgaactagct	ctgaacatat	ttcttcttct	attagagcgg	aaagggggag	263400
agattttcatc	atatccaaag	aagacttctt	gtctgataca	gaagttgttc	cacttcttat	263460
cttttcgttt	gcacagaaa	tgctcttgca	gcgtaaagag	aatccattca	aagctcccga	263520
tgctctgtcc	ttttcaagt	catttctttc	catcagtttc	aaaaatgctt	cattgttaaa	263580
gaaatcgta	tcttcttctg	catcaaatte	gtcacaattt	tcttctgccc	tggattgggc	263640
cggggaagaa	cttggtgccc	tcacttccct	tttattagaa	gcattaccac	agtcctcaac	263700
tttaatagca	cttgagacta	aattgttgtt	aactgacttg	acgttattac	tcttcattaa	263760
aattgcttca	tggcgtttat	tgacaatatt	ccgcagggtta	cttgggttgc	caaaaaagtc	263820
tctagagcaa	aacattttca	gtatctttat	aaaaacttca	gtgtatagtt	ttgcattatc	263880
ttcttggcca	aaattatctt	gttcgactga	cccttgggaag	aaggtaactc	tttcttcttc	263940
tttcttcgct	tcaagttcat	cgtaactttc	ttcttcttct	tccctcctct	ccccctcttc	264000
ttcttcggtt	tcgtctccct	cttcttcttc	accaccttcc	tcttcttctt	cttcttcttc	264060
cttcttggtc	ttgttgttct	tggtgttctt	ctgaggtttt	tcttctctac	tcgggttctt	264120
tattttcttc	tggatatggt	tagtgttatc	ttccactact	gcagaatttt	gcctgttatc	264180
tccaagtagt	ataacaaaca	ctcctccgaa	aggtatatct	ggcctgtgtt	tcgctatacg	264240
taaaacagca	tcaataaaca	ccaattccct	gcaactggcc	attgtatact	catcgataac	264300
aaggaacctt	aaatttgcca	acacttgtat	cactttccat	tttagtttat	cggccacaaa	264360
ttggattttt	tctagagtta	tttcatcttt	cctaaatata	actggaatct	caaatecttt	264420
gttgagggtta	gaaaaacac	tttctgggta	catcaaggaa	gaacctcctg	gtttcttgca	264480
tgacacacatg	agactctcct	ttgcagaatt	agatttgcaa	cacaggagca	tgtctgcaga	264540
tccagaatca	atcaagccgt	ggtcggttaag	aactgcaatt	agttccttga	ccattgtagt	264600
tttccagcgt	ccaggaagac	caccagatg	gaacacagat	attcctctct	cagcataact	264660
gacagaaatc	gagttgcaca	gatacctcac	tgccagagac	tggttgcaagg	aaacagtgtc	264720
caataccttt	ctggcctcct	tgacagtatc	tgggagcttg	accgcagaag	cagaggagga	264780
ggaggaagaa	gaagaccag	aaggggatgc	cgctgcagct	gcaactgcac	ttctgagcac	264840
ctgggttctg	gtagccagaa	aggggtatga	tttgatgttg	ataactggaa	ctggatttga	264900
attattgttg	ttgttctct	tttgtttttt	ggtaactggc	tcttggccgc	tggagcacgg	264960
acgtttactc	cccatgatgc	actctcggtg	aatgtcggac	acgggggtaa	aaccaatcct	265020
tttatactag	ttctacctcc	acgaggttgc	cagatggtac	gatactttat	tacaaaaaat	265080
ggatcgctcct	cttttgctta	agattttgtg	tgtaaccctt	atcgttacga	ttgttacaat	265140
atgtctattg	gcatttttta	tacatagcaa	ttcctcatct	tcttcaggcg	ccactgccgc	265200
cgctaccgtg	gcagcggcag	gagggatata	ttcatctgct	ttatttggta	cctacatagg	265260
gtccttggtg	cagtcagaat	ttttttcaga	agatatcgat	ttaccagaaa	aacaaaagaa	265320
aaaggctgat	aaggaaaagg	catcaaaaac	cgataacaat	aacaacaaca	acaagcgtca	265380
acgccgccag	acaagggaagc	agagcccgcg	cgccagaaac	aggactttaa	gaagggaag	265440

gaaattatcc	aaaagaaaaa	tacacgtcaa	gagaaagaag	actctaatta	atcggcgcag	265500
tctcagagga	ggaagaggat	taaagcgagc	atctagaggc	aaaggactgc	ttagagggat	265560
tggaaaaagg	acatttggat	tgtttttcct	ccccttttag	aaaggatggt	tctttttaatt	265620
atcatgttga	ttatatttagc	cccattaata	ttgttggggg	tgaatttctt	cgtatacata	265680
aagaagaata	actggttctc	tggggccatt	agtaatgaca	ctactattaa	ctccacaaa	265740
tctggaggag	taggaggatt	atgggataat	gtttctgatt	cagagtcagg	gacgtttcca	265800
tctagggtac	aatattccac	atcaacgaaa	aatttagtca	atatagccat	ggatggagat	265860
tatgcgacat	tagtaaggaa	cggaatgtcc	acaaatcaaa	ggccatatga	aacctacaaa	265920
gacgtggagg	atagccagtt	ctattttacat	tttttccacg	ttagaaattt	taaaccttta	265980
aacgggtgat	aaaataaaga	ccaccttgaa	agggatgaaa	gttttgtggt	gatcgaatca	266040
ccatattata	atggagggtt	tttatcatat	aatatcaaca	acccaaatcc	cattttacaat	266100
tctactgaaa	agccgtatat	taacacggag	ataacttcca	tcgtcagcac	cactggtaca	266160
gatgaaaggt	tcttctgcct	cgagaaggaa	tacgtcgaag	atggtgaaga	aggagttaca	266220
gaaaataggt	actttttacg	ccacatggca	agtaattatg	ttgtaaaggc	tagtttcaaa	266280
tcagtcatgc	ccactattga	aattagtgat	ctgactggca	catacaataa	taagaatagt	266340
gttaaaatta	agcctgtaac	tggtacttct	ataatttatg	aaaaattaca	aggaacaaa	266400
taaaatacat	tttaaagtat	tatgtttatt	gttttatttc	ttctaaaatt	ttatttttagg	266460
tatcctcttg	aatatggaac	gacgggtggg	gggttttttc	ttacccttct	ttgggtttttg	266520
gttagttttc	ttggccccac	catctggtgt	ttttggtcct	cctttccccc	ttccttttcc	266580
aaaggctttc	ttgaatgcac	cgcccactcc	tggtattgat	cccagcgcag	aagcgatcgt	266640
gccggccagt	gctgtgcccg	ccaggagaga	tgctaagaga	ggagccttga	ttattattac	266700
agcaataaca	atgacaatta	ttgttattaa	aagtaaacca	aaaagtatat	tttttctctc	266760
tcctaaagac	gccaccaatg	gcgaagccat	cttggctatc	ttattgttat	tattgttatt	266820
gttattatcc	ccttccatgg	ttgttgacac	taataatagt	gtgtattata	acattaaaag	266880
aaacatataa	attttggtct	aaggcacctt	caaggctgtg	aaaatattac	atgcagataa	266940
tggtgattct	gaaltcaggg	ttgtgatttt	tggtcgcacg	tttactgacg	tccaacatcg	267000
gcattatggat	agtacaaaat	gtcttgaatt	cctcatcttc	tttgcctctc	gtaccatcct	267060
gggggttttgc	tggaacaat	gtcctatatg	ttcccttttt	tatatagtaa	ttattcgggt	267120
ttgaatacaa	gtagtattgg	aaatggagat	gcgaaccact	aaaagcttgt	atattttatca	267180
aaaaattcaa	gtcaactttc	tcgtcatcca	ctcttctcat	ttttagggtta	atatctctcc	267240
cgtaaatgac	gtagataatt	gttgcatcca	aagcactgat	gtgcttttca	tgatattcgt	267300
ttaaaacaat	tttataatga	tttttaagg	gataactata	aagagaccct	tcttcttata	267360
cgatattttac	tgatggtcct	gctactacct	ttaaagtcac	gattgtgaac	cctctcaaaa	267420
gagctatcca	ttttttttca	tccgtaaaag	tatcgtcaac	cctccgcttc	cgcttaccct	267480
tcctccactc	cctcctcctc	ctcctcctcc	cccagctctc	tgccctttat	tatcgcttcc	267540
acctaagat	cttcttccac	cattatttga	aaatacaaaa	aaacctactg	caacaagagc	267600
agctattatc	atgaaaaatca	ccaggatcaa	aacaggggca	ttcgtattag	acatttcttg	267660
tacatgggtat	tggaatattt	tccacccttt	aatacataaa	aatttgggtta	caaatagtta	267720
aaaatatgta	accatgcagt	acaataaaag	gtgatcaact	ggtttttcat	ttctatatcg	267780
taatgaaaat	tttgcatcac	gcgtattctt	tgtaactata	tcggaatata	cttcatccca	267840
tacagcagat	gaacctcttt	tttcaggctt	aatttgcgata	tcctttttcag	attccatttc	267900
gtggaagatc	tttacaagac	aagattttatc	cgtttggatta	ttgaacaaaa	aatcgagccg	267960
ttttttccca	ttatctactt	caaaattctc	aattgttaga	gaaaaatccc	tcttccagac	268020
tacaatatct	tgggccttag	ccagtgttct	gattagcccg	tgattttattt	ttcctcccat	268080
agaaggagat	aaattgtagc	tggtttcata	cctcccttgg	tcatcacatt	ttactctaaa	268140
aattgccact	gtaaaatcac	tcagtgtgaag	attttcaatc	aattgaggaa	ttctcttccc	268200
tatatcaatt	tcaacgctag	ctggcgtttt	tagtactgcc	catgcaaacg	cattgaggct	268260
cttggttcggt	acaaaaaaag	ttaatgccga	tttcggatca	gggaaaacat	ccgcagccac	268320
ctcagttacc	caggcactaa	tggtgtcttc	agacattggt	aaagttgctg	ctgctttttc	268380
atattattagt	gctctaattt	tccctctggt	gccctgcatt	ttattattga	atacagtcct	268440
tccgtagggt	ttaatatgtc	ctccatacaa	atctacccca	gccatggaag	cagttgcagt	268500
atctgacgac	ctcgtgtcca	aaacatttgt	caccctagag	gaaatttccc	catctttaaa	268560
tagttctgta	agttctgcct	ccattattat	tatttttaat	gtgtacaata	aaaacaatta	268620
atttgaaaca	aaaggtaata	acaaaaataa	atacataaga	aagattcaaa	aagttttatt	268680
cattatcctt	atcacacaca	tattttttga	cggacaagta	cattaacatt	atcaacattc	268740
aacagaccac	taccatcatt	cattacaaca	acagaaaatg	gggaagtttt	gttgacacaag	268800
aacttccata	catcattctt	ctttacatga	atcaagaaag	tattccaatc	atccccattgc	268860
tcategagcc	tgtaccccaa	ccatgggttt	acaaatagac	tattaaattt	atcccatgaa	268920
cagtatttct	ggagagaaat	gaaggggtct	ttcccgtcca	ggtcattttg	tccaacattt	268980
tccactacct	ttacttcccc	gagtccaatt	agattgactg	tcgtcttttc	ttttgtctcg	269040
atcgcacact	cgccaatggc	agccaacaag	gaatcgggaa	gatcactcct	cattttcttt	269100

ggcgggtgaag	aaggagaagt	gacctgaaga	ggcgctggtc	tcttattatt	attttgttgc	269160
agttttatttc	cttcggcaat	tgcagtgttc	ttgagtttca	agaaatagtc	gataatgtaa	269220
aactcttcac	cctctcccag	ttcttcatac	atcacatcac	cagatgtcct	gatggcttct	269280
tccaccatag	ccaaagaagg	tgctccacat	tcaatcaaat	ccatcagcag	tttcttccct	269340
cctctcccca	ttgcgacgag	tttgttggcg	tacctgcacc	agggacattt	tctgttctgc	269400
ggcttgtccg	gaccaagagt	agacatcatg	tacttggatt	tgcagcaatc	acattctaca	269460
catacagcaa	ctggaatatc	aatattgccc	tcattcaata	gcttgtagag	aagacaagtc	269520
ttcccataca	aaatcccac	agccttcgct	gcagctgaaa	cagcaatttc	attgtcttta	269580
tctgttgggg	gaagattgag	agctccaatg	atatcactag	ccatctcctt	tactttgtca	269640
tcaattttct	tggcattcaa	aagattgagg	gttgtgggtg	atgcagtttt	ggcacgagga	269700
tctctggtca	atttattgtc	ccagtgatag	accattgggt	tgcccatcct	gtagtttccc	269760
ttgcatctca	agcacaattt	ttcatcatta	tcattttcca	tcttttagaga	ttttacacca	269820
catcttttgc	atatatatatt	caagtccttc	tctgataggt	caagtcgacg	tacagggaac	269880
gtctcatatt	tctccccaat	ctttgagtcg	ccagcctttg	tcaagattct	cagagccaaa	269940
aagtgttcta	gtttattggt	aaagtctctg	tcaactccta	tcttattggc	ccatgtaatt	270000
aagcacctca	aaaatactac	accggaaaaga	ttattctcga	tctctccatc	ggcagttcct	270060
tcaagaatag	agggcaaaat	aatatccaaa	tcagtctcgg	aactgctatt	ctgctcttgg	270120
taatactttg	tgaatgccat	ttttgcaata	gagaggatca	tgtttgaagt	gttggagggg	270180
acttcacagg	aggcagccca	agaggcggct	gcacacacag	atttgattgc	tgcataacct	270240
ccttctggca	gactcagacg	tacatcatac	aaagtctcca	atttttgaat	cacagtctcc	270300
aagtcgataa	acatcatctt	tgccacatct	tctcgtctct	caaatagata	gctggccaat	270360
tcttctcgaa	gcacttcctt	ctggtcatct	gatccacaaa	agactgggat	ggcaaaatat	270420
ctgttggagg	tttttgacgt	ctccaaaagc	acatcttcaa	catttctagt	aatctttaga	270480
catgatacca	atttcttgat	gaaatcaaa	cactgtccaa	actcgaacct	cttgtaattcg	270540
tccagtgata	ataatcctcc	tectgtatat	gaattggaga	aactaaatga	ctggtggaag	270600
attttgccag	cttcttccaa	ggcatcattg	ggcacataga	accacataaa	catggaggat	270660
tgttcagggt	tgagatctgt	attgttaatc	acttggagtg	catcatagct	agcagatttc	270720
atgtgctttc	cttcttttag	gaaggcatat	tctgccttgt	ttccgcattc	agtcccaga	270780
tcagaaagga	tggtagagac	cctcgagcac	actcctgcat	tattgcctgt	aacaaaggca	270840
taaatagccc	caatttcata	aggcacagca	ctgcaagcgt	acatcattga	ttgttccctg	270900
atcatcattt	tgaagagttt	ttcagcaaac	atcttggccc	agtgtacgag	agcatacttg	270960
cgagacaact	tggtttttct	tccttgatca	gatgtagcct	caacacttgc	catcaatttt	271020
gtacaacctt	tacaatcaca	gaggtttgta	tggtcaattc	caaaatcgct	caacgactgg	271080
aaacacgcgt	tcattgtgag	tgccagagta	gaggcggcgt	ttcctccccc	tgtaattgac	271140
tgggcccctt	gagtcactac	agccagtttg	ctgataccac	caccaccgct	agaagactgt	271200
tggtgttggt	gttgcgtgct	tggattaaag	gtttgcttag	agcaacaaaa	ggcactatca	271260
tatgcacaat	tcaagaacac	gctacgcttg	tggtgtttat	atcttccctc	tcctctgata	271320
actcttccaa	aatctaaagg	tgccatattt	aaaaccctag	acttgaaaga	ttcacgttga	271380
gcgtcgccat	acttttcaaa	tcctgcaggc	tctgatcttc	cttcactttc	aataaaaaa	271440
ttcaattctt	cttgggtgat	ttccctcaag	atgggggaag	cactaaattg	atcattcagg	271500
ttcttgaaga	gggaaataga	ttgcctcata	ataaaggctg	cgatattctt	gcgcacattg	271560
cacgatcttc	tgtaaaagga	ttcaaaaagt	tgccatctgt	tggtgtcgta	tccagaaaaa	271620
ttgatcacat	cttcttggca	caattcactc	tggttactgc	cactacatct	agagaaattg	271680
aggatgcgta	gccaacaatc	attctccact	agcgccttcc	tcatgtcaag	aatagtcaag	271740
ggtttttgtc	cccaagtctt	catgatgaat	tcttcttctg	gtgcagtttc	cgaatcgtac	271800
gagttggcac	ataattttcc	agtgtcacta	ttggcaatac	gcacatcctc	atcgggtgac	271860
acaaaaatga	ggatggggaa	tttttcatcc	ctggtaatgc	catgcatttt	cttgaatgca	271920
ggaatcatgt	tgtaaacac	gtacatcagg	gaagaatgaa	gagcttcact	gccaccgccc	271980
ccgcctccgc	acacttcggt	cttgcgcata	tgcttcacaa	aatctccccc	ttcccactct	272040
gaacagtctg	gtgcacgcac	aacagggtcca	cttcaaatgg	agtgtttgtc	gaaatcattg	272100
tagtaaaacta	tagcattttg	aatgcctgga	gcaatcacat	tgagtgtagc	catgatgcca	272160
gaacaactat	cagccacagc	cgattttcaca	cttcccattg	aaccactaat	atccatcaac	272220
gagatactta	gaccagttag	ttttgcagtt	ggagtatctg	tatcagcagc	agcagctttc	272280
ttagtccatc	catcttctaa	aagttcaatc	ttggaggaag	gtgcattttc	catcaaagag	272340
ttaaaaaaag	atgcagattt	cacagtagaa	ctagaactct	tgctcatggt	ggaagaagcg	272400
cttgaacgga	tacggcccaa	cattgtagta	tatctttatg	gtagaagttg	atgttgacgg	272460
taggaaacga	tctcgacaac	tgtatgctga	ggtgcctccg	caggccgctt	ttataccctt	272520
ttcgtgtgac	gtcaattgac	gctcgtacgt	ataaacaggg	tctattttta	gagatgacat	272580
ttttcaaaaa	cgtttatata	atggtgcaat	ttcctattac	tcaaaactagt	catatccggc	272640
taagacagga	tgtttttttg	gccaatacgt	ttctggcgca	atctgacatt	ggcccagccc	272700
agcgggtccac	ccttcgaaat	tgacatcagg	ccgaccacgc	ggtccacccc	ctaaactgga	272760

gtgagctgaa	aaatTTTTtca	aaagTTTTtgc	agatgaagaa	gaggggtgaaa	agtgggtacg	272820
tactatacac	tcagaggggtg	gcagatcggc	cctgttccag	aaatggctgt	ccagaaatct	272880
gggtcggaca	gattccagaa	acgttttctaa	ccattttctgg	aacagtcatt	tctggaaaagg	272940
gttacaattt	cttataactg	gtgcattatt	tctagcatat	attctagccc	ctttgtgcaa	273000
tctgacattg	ggtcgaccca	gcggtccacc	cttcgaactt	gacatcaggc	cgaccacgag	273060
gtccaccccc	taaactggag	tgagctcaaa	aaatTTTTtga	aaagTTTTtgc	aaacgaggat	273120
gaggggtgaaa	acacctgtta	gaaagtgggt	gctgggtcgg	agcatccgtt	caactgctgt	273180
ccagaaatgg	ctgtccagaa	atcagattcc	agaaacgtct	ataaccattt	ctggaacagt	273240
cattttctgga	aaggggttaca	atttctttaga	tgtggtacaa	taacttacca	tacattttctg	273300
gtatcttctgt	gcaatctgac	attgggcccga	cccagcgggtc	cacccttcga	acttgacatc	273360
aggccgaccc	agcgggtccac	cctctaaaact	ggagtgaagct	caaaaaattt	ttcaaaaagt	273420
tttgaatcga	ggatgagggg	gaaaacctgg	taacagaagg	catgtctacg	acggcttctg	273480
gttagatcga	gccggtccag	aaatggctgt	ccagaaatct	gggtcggcca	gattccagaa	273540
acgattctaa	ccattttctgg	aacagtcatt	tctggaaaaga	gttacaattt	cttagatttg	273600
gtacaataac	ttaccataca	tttctgggtat	cttcgtgcaa	tctgacattg	ggtcgaccca	273660
gcggtccacc	cttcgaactt	gacatcaggc	cgaccacgag	gtccaccccc	taaaactggag	273720
tgagctgaaa	aaatTTTTtca	aaagTTTTtgc	aatcgagtaa	gaggggtgaaa	acaccctctt	273780
gaaaggcacg	atctggatgg	tgtcgtccgt	tcaactgctgt	tccagaaatg	gctgtccaga	273840
aatctgggtc	ggccagattc	cagaaacgtt	tctaaccatt	tctggaacag	tcattttctgg	273900
aaagagttac	aattttcttag	atttgggtaca	ataacttacc	atacattttct	ggcaccttctg	273960
tgcaatctga	cattgggtcg	acccagcggg	ccacccttcg	aacttgacat	caggccgacc	274020
cagcgggtcca	ccctctaaac	tggagtgaagc	tcaaaaaatt	tttcaaaaagt	ttttgaatcg	274080
agaaagaggg	tgaaaacctg	gtaggagaag	gtacgcacgc	atactggcgg	cacatgcctc	274140
cagaaatggc	tgttccagaa	atctaggtcg	accagaatcc	agaaacgttt	atagccgttt	274200
ctggaacagc	cattttctgga	aagtgttata	ttttctcgca	tcttgtacat	taacttgatt	274260
atggtatagt	acacctctct	gcaatctgac	actgggtcga	cccagagggtc	caccctcgga	274320
acttgacatc	agcccgaccc	agcgggtccac	cctctaaact	cgagcgaccc	taaaaaattt	274380
ttcaaaaagt	tttgagacgg	aggaagaggg	tgaaaaccct	tgcaacagag	gggggtataaa	274440
gggaggtacc	ctcgacacac	ttagacacac	aacctcatca	ccctccgtcc	aatcaacatc	274500
atcaccccat	ctctaaataa	tccatcatgt	atatcttcgt	cgaagggttc	cccctcacag	274560
ggaagagttc	atggatgtcc	aagtgtgatg	atacaggatc	atgtggaatg	tctttctctca	274620
attttcttctg	tatgaacact	tctgactact	acaactggcc	tgccgaaatc	gggacgaac	274680
atctccagtt	aggtttcaga	gaaaccagag	tgggtggatgg	aatgtttgaa	cctgtcctaa	274740
agacctttgt	cgactcgtgg	aagaaagagc	aaggaaaaga	gagtttgaa	gaatatctgg	274800
actacaacgg	ccaagtcatt	gagatctaca	tcgcagaatg	gttgagacaa	aggccactag	274860
ccttccacgt	gtttacctat	acagatgaag	ctgtcaagag	tggattcttg	aacgaggagg	274920
atctagatat	ggatactgca	accaagtggg	tggctgaaat	tattagagag	aagaggggca	274980
atattcaaga	aataaaaagt	acccctagag	tagtcttcaa	tggcaatgtt	tgtagtgcac	275040
gtttctctaa	cactaagaga	aacttgtata	actttggaac	aaactataac	aatgttgtac	275100
attgtgattt	gttgtgccct	tttgcaaggc	ataggattgt	acattttctta	taaaaaataa	275160
aaacaatatat	aaagtctcagt	tgtattttta	ttgtctcaat	aagttactac	acccaattct	275220
ccccctctc	tattgagaga	tcccaatcac	tttctactgt	cactgtttgt	gctgttttgt	275280
gttctctctc	ttctctctct	tcttctctct	catcttctac	ttcttttagt	cgttcttctat	275340
tatatgcttt	aattatcctc	aatacattta	aaatgggtgc	ctcttctctg	ttaaacttct	275400
gattcttttcc	ctcaccacca	acatcgaccc	tacgtccatc	ctgggtatact	gtttcaccat	275460
caccagaaaa	aatttctaata	gacggatttt	ttatagaacg	tacaacaatg	tccactaatg	275520
attctgggtc	tgggtgctaac	tgatcatcgc	caataaaaaga	taaaacaact	ccagaaacac	275580
atggataata	ttgcacattg	aattgatcat	ataaagtatt	gaagatgata	gggtgaagatg	275640
ttggcccgta	ggcagacatg	atcaaacggg	gcctgaaacc	accctcatct	ctcattgtac	275700
aataattttga	acacatgaat	ctgattcctc	tttctgtacc	agaatatagg	gggtgtgtctt	275760
cattgaatct	gtttgtgagg	aaaaagtctc	gtgcagataa	tgcaaacacg	agacacgtac	275820
ttttgtctcc	tattcttctc	atttctgttc	catattcttc	ccttgtcttc	tcacgcgatag	275880
catccaattt	gttggtttcc	tctagttttt	cttctaggta	tttctctctt	gcggcgctcaa	275940
gggtatacctc	gtttggcata	tggatgacct	tattacgttc	atgcttaccg	tccattttttt	276000
cggccgaatg	tgtacaact	gcattcaaaa	ataatgccat	gcattgatca	aaatttttaa	276060
atgtggggta	catattactt	gaaaaaatct	tccctaacat	actcactgga	acaactaagc	276120
attgggtcatg	acaactaccc	actataggat	tttctatgcg	gaggcctttg	taggataatc	276180
gatgtgctat	atcgtgtaaa	ggtaccattg	attcgccccc	taaaatacta	cccatactta	276240
acagtttctt	ttcgactttg	tgtttgggta	tcccacaacc	catcaataaa	tcagcagtta	276300
caattatgta	cttgagtttg	tgaattgctg	cattgaaatg	tctatcagtg	agaattctat	276360
atagagggga	gaaagtgaaa	gacttgtcat	ggggtaaagc	ttctcttttc	actatattgt	276420



tgatattcct	caaatcttct	ggtagtcttt	tgagccaatc	agattcagat	tcagcacctt	276480
tccttgtata	gaaattcctc	atthgtgaat	gatgttgaag	aaataatgac	gttttttcca	276540
agcaggggtt	gcctgctgag	aggcgtatth	tattcacaat	ttcatccacc	tcattgttttg	276600
gagtattttt	atgcaaaag	acgatataata	caccttcagc	gagttcgctt	ctccgtatata	276660
ttttgataga	ctcgataatg	gacggtattc	ttctatagtt	tttaacgatt	cccttgatag	276720
caaaccggga	cgccatcttt	catgactcat	tttggagggt	gagatcatgc	aatgaaaaaa	276780
acttataagt	gtgtgtgtta	acgcacgtaa	ataccagcca	ttgacgtcat	atccgcctag	276840
tttttatcct	caacttctga	gaattgctca	acgtactcat	ctcttttccct	ggtacatgac	276900
gtcaatggaa	atgccgattg	acgtcatgtt	aggggtataa	aacgactggc	tcgagccctc	276960
ggattcacac	tttgtctgtc	ctgagcagta	gagaaggacc	cagtgaagaa	aatactcctc	277020
aacttccaac	aagaatgggc	tctgctaact	tgaacgaccg	aactctactc	gtggagggaag	277080
ggcaagtaaa	cgactacggg	tctaccaaca	agaaacaacc	agtgtgctgt	cacatttgct	277140
atctaggaag	aggaggaaag	attgcatgtg	tcacgcgcatc	tgttctgggc	tgcgtcctcc	277200
tgtctggtgac	tgtcatgact	ttgttgattg	tggtactggg	aactgcacca	gttaattgtg	277260
atgtgagccc	acagagctac	tcgccgcgcg	cgcagccgccc	ggtgcagttt	catccttacc	277320
attcttcttc	cacaaccacc	actacttcca	ctactactac	tactacacca	actccaccag	277380
atactaaaaa	agttgacgac	gactatgatg	acgacgtcaa	tattggaggg	caatcagtta	277440
ctgtgaataa	tggagggtgt	ttcatcaatg	gaagaaaact	ctcaaaagaa	gaagaaaaag	277500
caatgggtat	caatacagat	aatggaggat	ttgtttggaa	gaatgggtgt	ttttctcaat	277560
atggcaaaaa	ataaaatatg	gtatattaaa	tatgtgttgt	tttatttgtt	ctttaaaaaa	277620
taaaaaaaag	agtaactccg	tctacatagt	tttttttaaa	ttggaagata	taaatgtaat	277680
gagtaaaaag	tactgtgcca	ttacatgttg	agatcttgta	ttagcaacat	tagagaaatg	277740
ggggacaata	aagattatga	aacaagactg	attcaaagaa	ttaatgattt	ggaatcagaa	277800
attgaaaaata	aaactgaatt	atgtgaaaaa	atcaatgagc	agatgaaaaa	tacacaacta	277860
aaatatgata	aatgtttcgt	agaggaggag	acggaaaaat	tccgcaagat	ggaggaaaga	277920
gttaaatacc	tcaaagagca	gggaatccct	ctagaccacc	aagaaagacg	tacaattgtg	277980
gctgaaattg	acaagagtaa	caaagagtta	gatgcccttc	ttgaggaaaa	tgaacgtata	278040
ataaagctca	ttgatgaaga	gttggaaagt	atgaaataaa	ataaaagagt	aggagacttt	278100
tcagtttggtg	ttataaatat	atccctgcaa	ggagatacgg	tctcaaaaaa	tccatgactt	278160
gtcagtagcc	acggaaaaca	cacttgaaaa	atacaccaga	atagatgagc	agacaggaga	278220
gatgcataaaa	aaatatgata	aattatttga	agatgataaa	aggttccgag	aaatagagga	278280
acgaatcctt	acaaataaag	agaagggaaa	ccctctagac	ccagaagaaa	gacttgattt	278340
gtcggctgat	attgatagga	gtatgaaaga	gattgatgat	tgtctcgagg	aaataaacca	278400
tatagaatta	tccattgata	cattatttga	tgaattgtga	aacttgcatt	atggctctca	278460
aacaactaaa	taactcttta	tcgggcaagt	gactgtattt	gtacgtatgc	aattaggtcc	278520
atatttggtt	gcaataaaat	aatccacaca	aaacgtgttg	tttttctata	tcctataaat	278580
ctttggaacg	gaagttatga	ctcacataga	gtgtataaaa	ggaattggat	gttcctatgt	278640
ccgcatacct	tctaactgag	acatcgtttg	aagaggaaca	gagaagagag	gttaacaaca	278700
acaccagtaa	ggaaaatgat	acaatggaaa	aaaagactga	gacggctgca	acaacagaaa	278760
aagaccagaa	accgtctgtc	agtaaaaagt	ccagaaataa	agaaccctaa	acaacttcta	278820
ctgtttacac	ttctgtaaa	tgttaccttt	cttccataat	caagagtga	agtagtagaa	278880
gtaattgtcac	ctcaaccaaa	gaaaggtttg	aggagaagtg	taaataccgta	agcaagatga	278940
tggtcaaagg	ttcactgttt	ttgagggttag	tagtggacga	gtgtctgaga	cgttacaacc	279000
atctagaaga	cgaaatcgat	aaatggccag	atatgacgaa	ggataacttt	tacgtccaat	279060
tgttgaggaa	gggtttagac	aagaagaaat	tgaagaagg	atctacacat	cctgttgtag	279120
aagatgtttg	gaattccccc	atcgtccaag	aaacattcct	atcccagcaa	ggagaaggaa	279180
ataatcccat	aaagagacat	ctcatggatt	tcaataccat	cacctacgcc	gccaaaacaac	279240
taaaaaacttg	cttcgaacaa	aacctacgca	ccattttccg	gacacgacaa	cagaggggcca	279300
tatctggatg	gttagctgaa	aacgggttcg	ataaaaaagta	tacgaaactc	gtacaacact	279360
ggataatttg	atgtacctac	aagagtgtat	gggtggacag	tggtgatttg	gaaagggtaa	279420
aagaaggaaac	gaaaaatttc	gtgactcttc	ataggaaaca	tttatgtgtt	attagtgtata	279480
agaagaatgg	tacaatttcc	tattcacctg	aagagaaata	tccgatcccc	tcaatactaa	279540
attattacaa	gtttctacaa	acagagtatc	cacaaaacaa	gaaaatacag	aaaaatgatg	279600
ttgtcccaaa	acacaaacta	aagatacact	attgtacgtt	tgaccaaacg	accattcaag	279660
gaatttgttaa	agatttgga	gtgtggaagg	atatggaaga	acgacacaaa	caatcagaag	279720
atatacttta	caagcaaggga	tggtacctat	tattcgacgt	taaaaaagatt	aagaaattgc	279780
gtccaaaactg	gaactttcac	tctatccaga	cgcacggcga	aggcgtctct	gtactatttt	279840
ccagagaagt	ggaagaagta	gagactgttt	ccaagaaaaag	taagaaaaat	aaaaaaccta	279900
gaggagatga	ggataggaga	aattaccgcg	ccactaatgc	caagtacgta	gtgggtgtag	279960
atccgggaag	aactaatgtc	gtttcctgtt	cggtatttga	taccggtcaa	aaaagggtag	280020
tgagaaaaaca	cagaatgact	gcgaacaat	actatcaaga	atcttggtatg	acagatagaa	280080

gaaaggcaaa	cgaaacgtac	aagaagaaca	ataaagagta	caaagaggcg	ttagaggaaa	280140
taactaggta	cgataatggc	gaagaaatta	taaatgatgg	taacggtgat	acttctacac	280200
ccactaaaa	attcgaagct	tacttgaagg	tagtgaacga	gcactacagg	ttactgtgga	280260
acgaaaagg	aaagaaaaag	tacaggaaaa	atgccatgaa	agtatactct	agaaaaacaa	280320
agtgcataatc	taactttata	gatgaattaa	tccctaaaag	ggataaaatt	gaagattacc	280380
acattgtcttt	tggggatgcg	aaatttgcct	gcacgggaag	aggtgagcaa	tacgcatcac	280440
ctgccaggat	tttcgccaag	aagataaagg	aaagagtcgg	aggtgataag	aggtttactt	280500
tcgtggacga	gaaatatacg	tcaaaagtat	gccatcgctg	caatcaacct	ttaaatatgc	280560
tggagaagga	ttgtttttca	cgaataaga	aaagaaaacc	gccgacaata	gtaacaacca	280620
caacaacaac	aacaacagaa	gaagacgaag	aaaatggaaa	atggaagaag	gctacacctc	280680
tcagagaaaa	tagagatacc	agaagatgct	cgccgaaaaa	gacgcaattc	ggttacagtt	280740
caaaccgaaa	agtatcgaca	ggagatatct	ctatggaaac	gccagtacct	tcttccactt	280800
cctcttcctt	ttgtactcct	acttccatta	catgtgtctt	gggaggaaaa	ttcgtcgaca	280860
gggacttcaa	tgcaagcacc	aatattgttc	ataaatttct	agggtttttg	gataaaaagt	280920
taatggaaaa	gaaagacaag	atgccgttga	agtatcactt	tattcgagtt	gcctgagaat	280980
aggatgtaaa	aaactcggtg	cactagtgaa	atagtggcgt	agccctgaag	gatgagaagt	281040
atgctattttt	cgaccttacg	gtcattgtgg	tatatcaact	cttgcccgat	aaagggttat	281100
tcaacaacat	aaaaatggta	caatattcta	ggtgtgtttt	actattcttt	gcttgaaatc	281160
ccaaaaaggt	gacccttcaa	gttcacttga	tacactaact	tcaacaaatg	gggatatttc	281220
tgcccttcacc	ttttgcacaa	cagactctat	gaatttatta	cacatcaaat	tcaatttgta	281280
ggctgtttata	ggtccaatct	tcatatcggt	gtagaaatctc	gaacctcgga	cacgtgattt	281340
ttcacgggca	aggaaagcaa	tatcagtcag	tatagggaat	tcaaaagtga	ctttgttgaa	281400
tatctgcaaa	tactgaacg	tgcggtctgt	aatggtcatt	tcgttgtgcc	tgagaagtga	281460
tctcggacgg	aggcggttag	tcatgttttc	cagttgaggc	aaaagatata	ccatttcttt	281520
ataaattag	cgagaggcta	tttctttcct	cgcttccgtt	gcggcttttt	cttctctgat	281580
ttttatgtct	cggaccatgg	cttcggcaag	tctttctctc	tccattctgg	cgacttcgat	281640
tgatgaggcg	aggatctgca	catcaccatc	aacattttca	tcaatataga	cttcactctc	281700
atcgactctg	gggagcttac	tgcggcgggtg	tgtttctact	gattcttctt	cttcttctgt	281760
tccttcagtg	aggtcctccg	cctttctctt	actggtcggtg	ggagaaactt	cttttgcttt	281820
tgtctgtgct	gcagcttttg	cttccataga	cctggtggaa	accattttga	ggttcttgga	281880
tgatgttgta	tatgaacttc	ttagtgttca	aagagactga	acaagtgtgt	gctcaggtgg	281940
gatcctgcag	gccttatata	caatttcaag	atggggaaag	ggggggagaa	aataaggttg	282000
ccaactagta	cctttgtcac	ctacaccaga	aatccacaca	tgcagaaggt	ttgatgctag	282060
tttttgata	cattttctgg	gtccagttgt	gtgggggggg	ggtgggtataa	atatagccca	282120
caggaggaga	gcctgtgtca	ttctgtgtgt	atgtcttgta	ctgaacatta	ctgaaatggg	282180
gaatagtga	tctcgttcta	gtggtattga	aattgtacat	aaaaatgggtg	cacaaaaag	282240
gtctcataaa	actctttact	tgtcaaatag	aactgaaaga	catgcccgaga	tacaaaaagca	282300
gatcgaagag	ttacatcaca	aaacaaataa	gcaatttgaa	caggctcaaa	aggtcctcga	282360
caaaaacgaa	gagcgaaaga	agcatcagca	acagcaacaa	ataataattc	ctctagacct	282420
agaagaaaga	cgtgcaatat	tggctgaaat	cgataaacac	atgaaagaga	ttgatggttt	282480
catcgaggaa	agtgaacgtc	taggtttact	tgtagatgca	gaaatcaata	acttgggaaga	282540
aaaggagggt	gaagagggaac	atcttttgaa	acaaaaagaa	gactaaacaa	taaaacatct	282600
aaaaatacat	tatatgattt	tttgatttaa	taaaaacaag	tggaaataatt	tttaggtata	282660
ttttattttgt	tttttaccaa	caaggatcat	caatcaaact	atctattgta	acttcaatag	282720
aaggggttat	atcttctctc	actttctgta	ctacagcctg	tacaaaatta	tcacacttgt	282780
ggagcaactg	gttaactgta	agagatccaa	tctccatgtc	gccaaaactct	tttgggtggg	282840
gtgtatttacc	ccacctagaa	acggcttcac	ggttgatcag	gaacccaata	tcagatagta	282900
ctccaaaact	aaaattgaca	gtattttgta	cctcaaattc	atgctcatag	tgtgaaactg	282960
ccatgtcatg	ttcccttagg	gggtttttca	gtttgatgtt	tttggtagta	ttgtccagtt	283020
gcacaaaaag	attttccatt	ttctctgcag	tgtcgtcgtc	cactctgtct	tcatattctt	283080
cctcatcctc	ctcttctcct	tattctctct	cttccaccatc	gtcttctctt	tctttacatg	283140
ggagggagag	tgtagtgttt	tctttgtgtt	cttctctctt	ctcttctctt	attctggggc	283200
gcttggttag	tgacgagtc	tctaccagtt	cttcttcttc	catttttact	tccactttca	283260
cctttacttc	ttctggctcc	ctcttctctt	tgggtggggag	aagagtagag	gtggctgcta	283320
ctgcagatga	agatgttggt	gatgttctag	aagagaccat	tttgggggtt	atttttgga	283380
ctcgtgggta	gaacaagtag	tagtagtatt	agactcttca	ttgtttgact	tctttgttac	283440
tagtttgcga	aatacgggca	atgttttatt	atttatagat	aatgtcacat	atggcaacta	283500
cgcagataat	gttagatttg	atgaaaatcc	tggatctata	aaaaagaaat	caaaacactt	283560
aaacgaattc	gatttaaaaa	tttgcgaggt	taatatagca	aataatagta	gattaccaca	283620
aaaattttgtt	aatgctacaa	atagattaat	atacataaaa	cctccatctt	tacacgatgt	283680
ttgtaaagta	catggttata	taaaccacg	gaaattatta	ttagatgcaa	aagggaccca	283740

ggaaatcggt	tatatgttcc	gtcattttt	tggactactt	gtatatcacg	gatgggatgt	283800
cgatggttta	atgaatactt	tcggtttact	agaattttgtg	acacaaaatt	tcattaacca	283860
gttagaaatg	tacttagata	aagacccagg	tgaatatttg	tatcatgcag	aatatcttat	283920
gaatacttta	aggaatctca	gagaaattga	aaaagagcca	attagggatt	taggttggat	283980
taagtacagg	gctaattata	gaggtttcag	ttaataaagg	actgttttgt	ttttttttat	284040
tttttgtccc	aaacccattt	aaaataaatg	catggttagc	aaacttttat	gtctttacaa	284100
gccttgttca	aaaggactac	atattcattc	ccttcatttt	cttctctttt	ccttttctct	284160
cctttactgc	tggtgctgcc	gtttccatct	ccatcttctg	aagtttctcc	tccacttcct	284220
ccatctccac	gttcttcatt	atacatggaa	attgcccttg	acgtattgaa	aatcatgcgt	284280
cgttctcgcc	gcaaacgctt	ccgttcgtct	tcgggcaagt	tttcttccct	tttgagtaac	284340
tctttgtatt	ccttagatgt	cggatcaatc	acattgattt	cggcaccatt	ttgatatact	284400
tttccattat	ccatgaaaat	tttgatgtct	gttctctgta	cctctctgac	catttccatt	284460
agttcccttt	ctcctggagg	gactgcaaag	gttatactat	tgccactggg	gcgggagaaa	284520
agtctcgtct	tgaagtgggt	tcgtacactg	tcctaataac	atataggggc	atccgatcca	284580
gttatggaca	tgatcaaacg	gggcctacgt	ctcccacat	ctctcttttt	atattcagg	284640
gaagagacat	acctgattcc	ttcttctggt	gcagaataaa	aggaaccatt	tccgttgggt	284700
ctgcatagca	aggagaaatc	tcggtcagaa	aaagcaaaac	ctaagcaggc	tccaccatca	284760
tagcgggatg	tcgcaatcga	tgattttttg	tgacatccc	tcaaagaaat	cctttgttct	284820
tcagagacgg	cacaacgttt	cctcaattcc	ttcagttctt	cttccctcat	atttacttca	284880
ttctctgcaa	actttatttt	tcttctgaaa	ttatccatct	gtaaattatc	cctttcattt	284940
gcgtatatga	cggttgcatt	gaaaagagat	gacatataat	ttccaatact	tttagctgat	285000
gggtgtgtac	tgttgtagaa	aatcaaacct	agtttatcca	ttggaacaat	taggcacatg	285060
ttggagtaac	tgctgtgat	tgggtttgca	atgtaggccc	ctttgtatgt	taacctatca	285120
catatatcat	acactgatat	gcgggaagaa	aagccacatt	gttccaccaa	taattgttcc	285180
atgttacgaa	gtgttctagg	gcaatggctc	ggcgtgtagc	cacaccctac	taaaaagtca	285240
gtggttatta	tcatatagtc	acattttttc	acagctgcac	tagaaaagtct	ttcgtgtaaa	285300
agcctgtaga	gaggagaaaa	tttaaagttt	ttgtcactaa	gaagggcctt	ttctttaacg	285360
attttgttga	tattttgcat	atcaatgggg	agtgattcgt	tccacttttc	cccataatcc	285420
tcacacccat	catcttgcaa	taaactccct	aacctgggga	gattgtgagt	atactttaat	285480
ccgacttgaa	gagtattgtc	gcctgtcaat	tgtctcagtc	tatgcactcc	gttttttatg	285540
tgaagtactt	ccttatcttt	tagacggact	atatatagtc	cgtcccttaa	ttcacttttt	285600
gtaacccat	caacagattt	gaatgataat	agaatatggt	cgtatttgtc	agggttttca	285660
ccagaaatag	tgaatccatt	ttttaccttt	aaagtcttga	attcttccat	tttgatatac	285720
ttaaatgtgt	catactcggt	tgtcacagat	aataaaccta	cagaccttgc	cattgaactg	285780
tgacacagag	tgggcccagata	cgcaaatgtc	cacgttgctt	gaattactta	atctgtagat	285840
gaaatcggca	aaaaattacg	tctttagaat	gtctcggaat	gtcatatttc	tttgtgcgta	285900
gtacacaaaa	tgtggcactt	ttcaagccac	cacctctaca	tcccagtgag	gcggtggagg	285960
tggacgtcat	atttactgtg	aatatgtgta	ataatatagt	ccaattctta	aaaaatcaat	286020
agaccaataa	aacaacaaaa	aacattttgt	aatttattta	ttcaactaaa	atctcatcaa	286080
ataatcttct	cagcaataaa	cttacatttc	atatcacagt	gtacaatact	attcgtgcac	286140
ttaccaacac	ggtacacgta	attcttactc	gtcgagaaac	atgaactaca	aatattccct	286200
ttgtacacat	ttctggggat	aacttcaatc	tttacaacaa	acgctctctc	ctccttgaga	286260
acgtcaataa	accagttatg	tgcattcattc	aattcttgaa	taattggttc	cattttcttg	286320
acatcttctt	cctgttaaac	catcacatgg	aaaactggct	tgtatttgac	ccactcggcg	286380
acacatgctt	tgaattcttc	tccatacatc	tccaaatact	ccctcaatcc	aatttttgat	286440
tctggaaact	gttctgtgtc	gtgcctccac	atgcctgtga	aaaagttcaa	tgatgatcca	286500
aacattccat	ccacgagcct	gggacgttcc	tggtaatcgg	atgcacgcaa	atgctcttga	286560
atgggtccatg	ggaagatggg	taaatagtcc	ctataattgg	tgtacatgaa	gttgaggaaa	286620
gactgtttgc	cctttccggc	agttctcatg	ttatccaccc	atgaagtttt	cccagttagg	286680
ggagatcctt	cgacgaatac	aatcatgttg	aagggttaaga	cttgtctaac	tctctcgagc	286740
gagctctctc	tcttttatag	aggcaaaaaat	gtagttgccca	gccaccacct	taaaattcgt	286800
gcattataata	tataataatt	ttataattttg	tatcattcaa	cogtagtcca	tctaagagtc	286860
aacatgtcct	cctcctcgtc	gtcgtcggtc	tcgttcgcga	tctccaccta	ccagaccttt	286920
ctcaaggctc	tggctcaccc	agatctggtg	gacaagatta	cccagaaatg	tgacgagaca	286980
ggaagaaacc	agaagtgtcc	catccagttt	ctggccgaca	tctcgcacct	gatccaagga	287040
gaaagaaatg	gaggaaatct	gttccctttg	caccggttca	agaaccaacc	acatctggaa	287100
ccaagaatag	tgggaagtct	tcacggggaga	acattggaca	atgacattga	agaatcatac	287160
tgattttttg	tcaggatctc	gtataatgga	gtattttcct	atgtgaacgg	cgtcaaggag	287220
ttacagggcg	ttctggacaa	gaaaaatatct	ggatctggat	ctggagaatc	ctcctcttct	287280
agagctcccc	tgatcccaat	aactgatgtg	gatttgcgtg	acattttcgg	tactttggta	287340
gttcttcccc	ccagatctaa	agcgtaccga	gtcatcactg	aagctgttct	agcactcccc	287400

ttcaatgaat	tcagtaacaa	ctggcctcct	acaaatatca	aaggagcata	cgtgtctaga	287460
gatttcagga	tgtttaatct	gttgcccggt	ttagatcata	tagaaggaga	agtggggga	287520
gaaagtgaat	gggaatccat	acacgcatct	gtcgtcaagc	gaatggtcac	cattatgcgc	287580
aacaaagctg	agaagaaacc	tccatcaaca	tctagaattt	ttagagtgtg	tgtggctgaa	287640
ccagttaatg	atgcagtgac	aaagatccct	atacgtgtac	tcagtaaatt	attcggttca	287700
agactcgccg	gtattctcca	gaaagtgtac	tcttattcaa	tgctaaatct	tccatatctc	287760
ctgtcttcaa	attctataga	catcaagcaa	ggagtaaaag	gaattacatt	atctatacca	287820
tctgcaagaa	aattgggatt	ctatttactc	caaaaggata	caacattaca	atcttcccta	287880
tcacaagatg	ttgccgactg	catagtttca	atcaacgctg	gtattatttg	tgatgatttc	287940
tctgaaaaaa	tacgacagtg	cattgaggag	aaaaacaagc	cagaaaactg	ttgtatgtgc	288000
ttttgtgaaa	ttgacaagac	gcccgaattt	tcttatagtg	aacatgtggc	aaggcacaat	288060
ttcttccccg	tccacgcatt	ctcctcatca	catgatgaca	agtgttgttg	agcaaagatt	288120
tgttccgaat	gtatattccc	gtacatcatt	tccctgtatg	agaaaatgac	tgggtgggca	288180
ggtgtaaaag	ttgtagattt	gttccagtg	cctgggttga	aaagtggcat	gctcaatcta	288240
aagggaagat	gttacgagtt	ttcaaatttg	tgtaagagaa	tgatactacc	atacacatcg	288300
actcatgtgt	cttctctctt	tgatgctacc	ataaatcgcg	cagaggcttg	tttttactcc	288360
ctagagtttc	tccagtatga	ttttgaaact	gcgaggagaa	ttgcgcatgg	agctaaagac	288420
attccccatg	tttacaataa	ggtagttaaag	aatgtaaaag	atttgatag	actgtgtgcc	288480
ttgtactgtt	acaaatgcgt	atctcctgtc	gtatgtgatg	agccgaatga	aagtacggac	288540
tatgaaatgg	tggatgtaac	tccccctcta	attaatctta	ccgagattgt	tgattcggaa	288600
gagtatgatg	atgggtcccg	aaatcatatg	tggccagcaa	aatcttactg	taactttatt	288660
gcgggttcca	gtggagaaac	acccaccatt	agtagctgca	gagatgctgt	aacttttcta	288720
ggaagagcac	cgaggaagaa	aatggcagga	tgggatgatc	aatcggcagt	ggggcaagcc	288780
attatagcgc	tagccaactg	gagaaagagt	ggggaattgc	ccaaaaatat	gtttgattta	288840
ctagaagggg	taaatgccgt	actttataga	ggcgacagtt	tcttgttacg	tgcgataaac	288900
taccctgtgt	ttattggtag	atccatgagc	cctagtctgg	aactcgttaa	aagaaagggt	288960
aataaaaattg	ctttaataaa	ggccttcttc	acagagaaaa	gggtgctgcc	agacgcattc	289020
aaaaagttac	ttgaatgggc	agaactatta	gtcaaaaagt	atctcatgga	agttttactt	289080
cagacgccag	aatgtgtcat	acaccgcgcc	cattcatttg	taggcaaaac	tctcctcatt	289140
actgacgaat	tgggttcacat	gcgtccagat	gatgccacaa	gaaacgccta	tatccagaac	289200
ctaaatgcgg	ccagacagaa	tgcggctgct	gcagcctcat	tttctgggtc	gctcccaaaa	289260
cctgaatttg	tcccctgcaa	agaaaggacg	attgaaaggga	tgtatgaaaa	ggacaattgat	289320
gatgttagag	ttgtaaattg	tccttcatgt	aaaaaggcta	tccagaaata	tggaggttgt	289380
gtgaatgtgt	tttgtgaatg	tggaaacaaac	atgtgctgga	tatgtgaaga	gaaggtttct	289440
cctgctgatt	ctaatacattg	tgtggagaaa	cacaggattg	tttatagtaa	ctgtgttagg	289500
gttaaataatg	ccttagaaaag	tatgtacggg	tttgagattt	gtaccatgaa	aaatgtagaa	289560
gaaggagtta	aaaattatta	tgtaatggag	aattggatttt	tctttgatgt	acaagaaatg	289620
gttgctaaga	aataataaaa	cacttgtgaa	atgttataac	ttgttttatt	gtccattata	289680
ttcaacaaac	tttaatttct	ttgcatgcct	tgttcaaaaag	ggcaacataa	tcgccttctt	289740
tctcctccct	cttctctctt	gttctctccac	tgcttgtggc	tacttctctt	tcttccccac	289800
gttctctcgt	gtacgtctca	attgccttaa	ctgtgttgaa	aatcatgcga	cgctctcggc	289860
gcagacgctt	cctttctctc	tctggtaatg	tttcttctct	ttttattatg	tccttgcctt	289920
ctttagaaga	gggttcaatc	acttttattt	ctacaccatc	ttgatatact	tttccattat	289980
ccatgaagat	tttgatgtct	gatccagtaa	cctctctgac	catctccatc	aacaaccttt	290040
cattcggagg	atcgatgaat	gatatttcat	taccctttcc	ggatgcaatg	aatgggttat	290100
taaaatgggt	tcgtatacta	tctctgatgc	atataggtgc	atctgagcca	gttatggaca	290160
tgaccaatct	gggcctacgt	tcacccacgt	cctctttctt	gtagtcgtcc	gaagagacgt	290220
atctgattcc	ttcttctgtg	gcagagttaa	aggaaccatt	tccattgggt	ctgcacaaca	290280
aggagaaatc	tcggctcagaa	aaggcaagaa	ccagacaggc	accgccgtca	tatctggatg	290340
atgcaattga	ggacttttta	tgcacatctc	gaagagaaat	cctttgttct	tctgaggcag	290400
catatagctt	tttggtattct	tttaagttttt	cttccagaac	acttctttcg	ctaactagaa	290460
ggttgacttc	attttttgca	aatctgattt	cctttttcgaa	attatctaata	ttttgattac	290520
tcttttctag	actgcagcat	ctgacgatca	tgaagaaggca	tgacatgtaa	tttccaatgc	290580
ttttgactga	cgggtgtgtg	ctgttgtaga	aaagtaatcc	aagtttatcc	attgggtacaa	290640
ttaggcacat	atttgaaaaa	ctgccaacta	tgggggtttgc	gatacttaag	ccattgtacg	290700
ttaacctatc	acagatatca	tacactgata	cacgagaaga	ggtaccaccg	tgttgactca	290760
acacttggtc	catattctta	agagtacggg	tacaatttct	aggagaaaaa	ccacacccta	290820
ctaagaagtc	tgttgtgatt	atcatataat	cacatttctt	aacggccgca	tttgaaaagac	290880
gttcatgtat	aggggtgaaa	attttaaagt	cttgctcatta	cttgctcatta	actagagcct	290940
tttcttccac	aattttgttc	atttctctgga	atgggcttgg	tagtgattcg	ttccattttt	291000
ttccataatt	agggtcttca	tcttctaaga	aactaccttg	ttttgggaga	ctggatgtgt	291060

atcttagtcc	aatctgaagc	gtaggatctc	ctgttaattg	tctcaatctt	tcgttaccgt	291120
tcttgatgtg	gagtacttcc	ttgtctctaa	gacgaactat	aaaatgtcct	tccttttagct	291180
cacttctcct	gacactcttg	actgatttga	atgttaatag	taccctgtca	taattttcgg	291240
tattttcccc	agacaaacaa	ataccatttt	ttacttttaa	ttgtttaaag	tcttccattt	291300
ttataccgtg	taaccctaac	aatgggtgtg	cagtataata	aatctacgga	ccatgctctt	291360
gatctgggaa	gtgtactcgt	gatttctgtg	gaaattgtct	tttcctattc	attatttcca	291420
ttacagggaa	agatgataaa	atgaagacgt	gcgatgagat	atgtgttttt	tcaatatcgt	291480
atatcaaatt	tgggcaaatt	acataatttt	tagcgtcaac	actcatcttc	atgcaagact	291540
ttttggagt	aaattactaa	aaaaggctat	atctcatgtg	caacataaaa	cacatgtgca	291600
gacattaaat	tggtttgtct	acccccccaa	catcgttttc	atcaactctt	tcctcacctc	291660
tatataaaac	cagcctgcac	cccctcggaa	cccacagtcg	acttgtacc	tcctgcgaga	291720
tgacatctcc	agctccatca	ccctcttcca	cccccaaatc	cagttgtacc	actattgtaa	291780
accgatgtgg	tttctctctt	gacaacaaca	aggaagtggg	catctacgac	accaattcca	291840
aattcaagt	tgaacccaaa	aatctggaac	taattgggtg	actttctgga	gtctctgata	291900
atgttggtac	ccagatatcc	cccgaccaga	tatttgtggg	aacatataatg	gtcaaatata	291960
actgggtctaa	atctgggtcat	gaacgcttca	gtgacatgag	taacaactgt	ctgggacaata	292020
ttacacgccc	ttcagaagt	attgaaagt	tgataaagaa	aacgtccagc	gactttaaaa	292080
tgaagtacac	acgttctctg	atggaccaca	ccgagaaata	ctatttttct	ggtgacccaa	292140
aattgagcaa	aattagtagt	tgggtgtacaa	cccctatacg	acagtgggta	tgcaactccg	292200
tctagaaaac	tttgtcaagg	aagaacatga	aactgtcgtg	gtgcacaacc	cttctggaat	292260
gactggatct	aacatattta	atagttcccc	cgtgtatttt	gaagtgcaca	atgagatgga	292320
cgccctaatt	tttatggcgg	cttctctgaa	gcacaatagt	ttatggggag	aaattaacgc	292380
caatatggac	ttgtacacgt	ttgattatgc	gggtgctttt	ctggacgaaa	gatggtgcca	292440
ccacgagaag	agttttctct	tcgtccgagc	acaacttatc	aactcgtatt	acaagtgcag	292500
gagaaaaatc	atgcaagccc	tggacaataa	ctacaacaac	agaataaaga	agaggaagaa	292560
tgttggtgga	gcacctgcgt	tcacatttat	gagcggggac	ggagagggag	gaaaggaagc	292620
cctagaagct	agtttcgatg	tgattggggg	aacaagagga	ggaagatttg	gtgttgattc	292680
aacaccatgc	ccccattctt	cagccatgca	actaaaactg	gacaatgaag	gaaactatgg	292740
atgtattgcc	tgcttcgcat	caatgttctt	tgtattggag	aaccaggtg	atgaatcttc	292800
cttcatatca	acggatgcct	ctaaaattgg	acaagcgcaa	gcatggatag	atgaacgact	292860
acgaaaacaat	gaaaaatggg	gagaagaaaa	taatgtcttt	aaaaagacct	tccatatgct	292920
ggctgatatt	acccaaaagg	ctcatgaaac	tgcctatttc	aataccatcc	cacttggacc	292980
caatggcagg	cagtgggaatt	ggcctactca	cactgtggaa	cctattgccc	atgaatttgt	293040
taccatttct	ctagtaaaca	cattgaaaaa	tctaggggat	agaaaacttc	cccgattcaa	293100
ttttgatatc	ttgtacaact	tgcctaatcc	atttggaaaa	atgttgctag	tgtttattca	293160
aaattgtcac	attttaactg	gacataaaaa	caatgaaaat	gtgggtgcctc	gaggttctgc	293220
ttctgggaag	tggtggacta	tttaatttgt	gggtgtgaac	atgtggactt	ttcaagtaac	293280
aaaaatgaaa	gttgaaaagg	atagaaaaat	atccgatttg	gcctgtatgg	aaactctccc	293340
tcgtctacct	aatccaggaa	gcactaccgt	cgatgacaga	atagttttta	agggattctg	293400
tagaggggaa	aatctaggga	gtgtaggtga	agtcgtatcc	gacattacac	agagtgtcaa	293460
gaatttttgt	ctcatggttg	aaaataggaa	atcttagtgt	gataaagaaa	ctggtttcat	293520
ctcttcagaa	tcgatagtgt	ctgatccctt	cttttacta	gaagtgactg	gctgtagatc	293580
taatcgtgcc	caagatacta	ttaataatgg	ccgagttagt	gctcgtgtaa	tgaggatcct	293640
aaagtcacgt	gaaggtgctc	gtgtatggtt	ggccaaggat	gaaaatgcca	tcacttttga	293700
aaacgttaac	cacgatacgg	ccatctctac	ggacgctatg	gagcgagcta	tagggcagca	293760
caagatactg	tactatgata	ttgaaacaac	agataaagat	ttcaccgaca	aaaaatcagt	293820
catcacatct	attgggttct	gtttgtgtac	gggaggcgat	atgacacatg	gaggagagag	293880
aggagtattt	ggactgggtg	cacctggatc	cgacgtggaa	aaggtgaaa	agactataat	293940
aaattcgtac	gattcctgaag	aaaaggaaga	cattatgaaa	cagtgccttc	aagtgattga	294000
aattttcacc	aacgagtttg	aatgtttgct	tggttttgga	aagtatatag	ataaagtga	294060
gcctcacgtg	attagtgggt	ggaacaatgt	agcttttgac	gacccctttg	tctttactcg	294120
tatcgtcaaa	catttgagt	atcacacca	agacatgtct	tattgtgtag	cagatgcac	294180
tagacgagaa	ctgtctcttc	ctagagcaac	agaaggagga	ggaggaggag	aaactccata	294240
tcagatgaag	acccctcaag	aaagaataca	actagcaagc	actggatttt	tcaataaatt	294300
gggaaaattt	gtagacaaga	aaactggcat	gttgaaacct	gaaatgactg	cagattttat	294360
ggccggggca	gaaagtcagg	ccaataccaa	gtttaaggaa	cgcaacaagt	tatcctccag	294420
taataaagga	tcagcaggat	ggttccagaa	aattattggc	ggtatgtgca	gtgctattcg	294480
ggtggatctc	atgaaagtgt	gcgaaaaggc	ctataaagaa	tcctctctctg	aatttaattt	294540
gaacgcgtg	ctcgccaaag	tgagtagtgt	cgccgacaag	gttaaaaaatg	taaaagatga	294600
agtagacct	cactttcatc	tattgggatt	cttgaagctg	aagaaggccc	aggatcaggc	294660
aaaagtacac	gtctattgtt	gcaaggatgc	ctacttgact	ggtatagttt	ctacctccat	294720

caacaaggaa ggggagattt ttaggctgtg tatggactct gctttaaccg aggcggctcgt 294780  
 gacagccaac ctggccactc ctctatgtat aggagaagga gcaatctgta gaaatatggg 294840  
 agaagaaagg gcagatagaa gaggtgtggg agtaagaaga cactctattg ccacagacac 294900  
 aaagggaagg atggtgagtc aacctatcgt caatcatgtt ccctatcaaa cgattgacat 294960  
 gacaagtttg taccggatga ccatgtgtca gaataatctg tgcaccacta cctttgtgac 295020  
 ccatcgacaa attatgcaac tgagggatag attggtactt gaaaaaatga aaaacaaaac 295080  
 caccgactct ttattgttgt tggacgttat tgacgagtgc aatcagattg tgttgtccga 295140  
 gtacagaccc attgatattg cagtcgcac atggaagaat agcaactcta atagacaaac 295200  
 tccaattact cgcatagagg aaagtgtggg tctaagattc atagaaaatt tggatgccga 295260  
 gaagacaaat aataaaacgt ggtgcaccaa tacatctccc aatatgaatg tcactgccgc 295320  
 aggtatggat tacttcccc agattgtgtg tgacattaat atgcagtttg cggccaaggt 295380  
 gaatgatgat atgcatatag cccagcaag tttagagtat atgcttcaag tattgcccat 295440  
 aatgttaatc gacagaccgt acattggcgc acacataaca gctggaaaat gtcgtacatt 295500  
 ggaggatatt ctctcagaac tcgaaaagga cttttctgtt gaaaaagatg aggaaattat 295560  
 aagaacccat tggacattta aggttcaaaa acaatacgtat ttctgtcata gtcctgtgac 295620  
 ccaaatggct cgtcacatta ttgaatctac ggaagaaat atccgtgatt atgaaggcaa 295680  
 tgaaaaattc gagaggttgg ttagcttatac agacagaatt tatcgccgcg ttggcgcatt 295740  
 tgattcggcc aatgatccag ctgttagact gtggtcttct cgtctaatac atgttggaat 295800  
 gttgggttag acatggaacg taaaaactga cattcttaag ggaatcatcc ctcaaagca 295860  
 agccacttac agagccgatc gagttgtgat gcagaacaag gccaaaggat ttgccaagat 295920  
 gggagacatg aaaagagctg gtctaataca agttggacaa aatattatga agctcggat 295980  
 gaattccatg tatggtcatt tggccctaag agccgtaag agtttgcgtc 296040  
 tggatctgcc aatactgcct caagtatttc caacatgtca gccaccggag gaattggagg 296100  
 aggcacaagg cactcgggtga cggccaatca gattacagaa aatgctcgat gtgtatttgg 296160  
 caatattggt tgtggattac agatggctct tcctggtact aagcagacgt acggggatag 296220  
 agattctgta ttctgcgtgc ataatttgt aggtgatgga ggaatgatac cagaatatga 296280  
 tgaacaaact ggcaaatatt attatgtgat ggatattgct ctaaaaataa aattggctgc 296340  
 aattattccc atcctagtca actcgttaac aaagggcctc cagttttag agcgccgaga 296400  
 cgctggtgtg ggcattgatga atatcgccca tgaacgtcta gctgtcgtc gtcttttgtt 296460  
 tgccaagaaa acataccata tgcttcactt taagtttgca tcctttatca agagacctag 296520  
 gataaaattg aaatcaaccg ataacaataa tcctttcattg attcttagag cggccgaagg 296580  
 ccacgcagat gggatgtgtg tccccataa ggaagaggaa ggaattcatg acgagaagag 296640  
 aaccgctggt aagaaattga aaagctttt gtggatggcc atggatgctt ctgttatcaa 296700  
 tatggaggaa tggtttacct cttcacctac gtggaagggt aactggattg acgcatgac 296760  
 caacttgtat gcctcacaac ttgtaggggt gatggaggcg gtgacgcaag cgaatgcagc 296820  
 ttcccgcccc atagaagcgg gtacagaaat taagtttga tcccttatca agagacctag 296880  
 ttccacccct taaaaaagg gagcctttgt gaagaaggga attacacca ccaccaact 296940  
 aaagggtctc caatcattga ttgcaagatt ttaccaaaa atagaggaaa agaaaagctg 297000  
 ttatttggat gtgatgaaga atcatgtgga gaattttgca tcccatataa caaatcctgc 297060  
 tatgatgatt actagtccc gagtcaacaa gtttgatacg tcaaaaagaac agagtagacc 297120  
 taatcctcta gctctagcga taaataacca cctgaaccct tcttcagaaa tttcattggg 297180  
 gcagaaattt aagacggtga catcagtttc ttcttggagt ctttcggcag aggaaggga 297240  
 agtccctgct ggttatttta acgctggtag cgttcggttg gatgccacca acatgaaggg 297300  
 aagtgttctt gcattttcag tcaagaattt atctgttgtg cccaacgcca tcacatctgt 297360  
 atacaagatg gttgagagcg ataagacggc aataaaatcc atgattgcta aaaatgtaga 297420  
 agtgttgtgt tctacatctg ccaatactgg attttctctg agaagaggag cattgtcatt 297480  
 taatacaggc gtcattgtta ttgaggggg ggcattgggt tgtatacga ctctaaataa 297540  
 taaacaaatg ttattgtttg ttgaggggg aaaggattac ggtgaagacg acgacgacga 297600  
 cgacgaagaa gcagaagaag aggcagaaga aaatggtgaa aacgaagaga acaaagggtga 297660  
 ctgtgtcacg gaaaaaaga tccctggacg aagcactaac aaggatgttg gtgaagaaac 297720  
 taaaacaagc gagaaaacg agggagaaag aaagggtctt aagacggcaa agggaaagac 297780  
 ggaggaaatt gctagtctgt tgagtaaatg tgggaagaaa gatgcgagag atgtcattct 297840  
 tgaccgttta ctaaaagcaa cacattcttc ttgcaccaac aatgaagaga gaaccagagt 297900  
 cttacaacaa ttagcaatt gtacattatc ttctatata acttcagtca tgaaattgga 297960  
 ccaaagagta gcagaccaa tggaaaattt aatatctcaa ttggatcaa tacgtaatt 298020  
 ctccaacaag aagaggcaag aaaaggagg gccttttaag tctgaattgg acgcatggg 298080  
 tgctgcagtt aaggtttaagt ttttcccgat ttagatgctg tctagaaaat tgactcaaga 298140  
 ccattggaaa aagtgtcccc tgtccatccc agaaacgcgt gaagaaaaac cattaatggg 298200  
 tgtgctttt gaagttgcac tcaattctct aataggaaaa cacaagtgca cagatacatg 298260  
 cgacatggct tgttgtcaat cattgtattt tgtcctcttg tacacgctag ctttaaaatt 298320  
 tgagaacgaa agattggccc ggcaaatgg cctagatgac tctgtagatt tgatggctga 298380

gatgttggtc	ggaggggata	aactattggc	ccaggaagtg	ttaaaaaggg	taaaagatgc	298440
tcaagataga	aagttggtga	aatctttatt	gcctttaaat	tataaccatg	acacaaatag	298500
aatttatatt	ttgtttgagt	ctttaagggt	tgctcagaaa	cctgtagctg	gtatgagtgt	298560
tagtgaaata	aaagacgctg	ttagagggtc	ggccttttct	accactacag	gtactgtgtg	298620
gaattatact	gatgaaagat	tttttgacc	attgtataac	atggatgaac	tttgtaacga	298680
acgtgtcaat	ggaaattgta	aattgtcctt	tataactggg	atttatcata	cggcagcagt	298740
agaattggct	gctgcatgtc	tatcttgtgt	tttgtaagaa	atgataataa	aaaacatgta	298800
taatttcatt	ctggttttta	ttttatatcc	tcaagagttg	tagtagtagc	agcagcagca	298860
gcagcaatag	cagcattatt	ttcagcgaca	gcttcttcat	ttctattgat	tgatatctgg	298920
tcattttcta	cagcccatgg	tcgtttaagg	aatattcttt	gaagcagaaa	cacataccca	298980
gcaatgattg	caatgatcaa	tattaatagt	agcgaattaa	ttaaaataag	cattttgtcc	299040
acattctggc	taaaataaaa	aagtttatct	ctcaatgaaa	aattttttatt	gaccaatcac	299100
ttcgggaagt	tttgagcat	gacacactat	tactccaaaa	ctatgctcac	tcgtatctag	299160
tcccacaaca	tcacgcgtca	tatttaaaac	atgcaccttt	aaagcctggt	ttttcttata	299220
caattctgta	ttgattatag	ggtttacaac	ctttaggcct	tccttttagag	atactatttt	299280
catgtatttt	acaatcgaa	atgagttgtc	ttgtgcacca	gacgggagcc	tgctcatgaa	299340
ctctacttct	ctccacgctg	tattgagcga	gagatcatta	agcacaactt	ctgaagaaac	299400
gagaggttgt	aaaatagagt	tatatattag	tttaattagg	aaacatttgt	ggaaacttga	299460
actgttaata	tcaatagaaa	ctgagctata	tgctgacgac	atggccttgca	ttttcaaatc	299520
atgttcaaat	cttactacat	aacccaagaa	attataatct	gcttcttcac	catcatcttc	299580
ttccatggca	tcctctcctc	ttggctcttt	tagtacaatt	tccttttggt	ggttagctcat	299640
gttgatattg	ttcattgagc	tatgatttcc	cacaacatca	tcaaagaaaa	ccccttctgg	299700
gataatgcat	gcaccgacag	aggtgtagaa	taatcctcca	tatctaataa	caccgactgc	299760
ggtagcggtc	tcttcatcat	agtaaccgag	cttgattttc	acatgactca	tgggtccgat	299820
agagatatgt	gaagccggtt	cagtagcgga	aaaaatagac	acttggttga	acattttgtc	299880
gtcaggaaga	gtctcgaata	aaataccgac	aggggaatta	tcttcaacac	tattaaagt	299940
tgttgcataa	gatagccgcc	ctgaactact	tgtattataa	attggtatat	cggatgctat	300000
ttcagaacaa	agatctccta	taaaattgtg	attagggcag	accaacatcc	cgtctccagt	300060
gtaagacgat	acatgaagtc	ctgggccagt	tggcacgaaa	atggcagccg	ttcgtgaagg	300120
atttctcacg	tcaactcgtc	tcaccttcac	tggtactttt	ccaaatcgtc	gtggatctac	300180
attgaaacta	aaatctcttt	cgtgggggat	acttccatcc	tttaaagaga	agagaggcac	300240
tctgtaaaaa	ggagaccaat	ctgaatagta	ccttaaaagt	gttctatttt	tttcttctag	300300
agtgggtacc	gctctgcggt	gcattcggta	cacctccaaa	tccgtttcct	ttacttgaat	300360
cgattttgct	ccttcatttt	cccagagcgc	atcagaatcg	tacctgtcgt	cgaacgaatc	300420
ttcacctccc	attttatttta	ttttcctcta	tatgataaaa	accacacctt	aaatgtctaa	300480
tattttacta	caactgaaaa	atggcggtaca	ttgaccaagg	ggcgttttga	gccaaatctg	300540
tgatgcagca	acaatacctt	cttccaccct	ctaataaggct	ctccaaacaa	caacctccac	300600
tggcatcatt	ttctcttcaa	ccttctctct	ctaataaacc	taggagtaca	tcaagagtaa	300660
cagacatatt	tggtgtgata	acttggtgtg	ttttaattgc	tgcttttata	agcaactcat	300720
ttagtgttac	aaaaaatgtg	gtaaaacttt	ctaaagaaca	aaccgagaaa	atactagaaa	300780
aagatttgcc	tgataaggtg	tacaaattgt	ttgaaaattt	aaaggatgga	acatttggtg	300840
taggggaaga	tgaggaggaa	gaaagggagg	agagggagga	aggggaggaa	gagcttgaaa	300900
caaaaacat	aagacttaaa	agaaaatgga	agaaaatgac	tgaacaaggg	gatcaaggaa	300960
taaaagtaag	gaaattacat	ggcccgagag	gagaaagagg	agaaactggt	ccagcaggag	301020
cagttggccc	tgcaggccct	caaggagaaa	gaggagcaat	tggaccggca	ggaaaggatg	301080
gagcagttgg	ccctgcaggc	cctcaaggag	aaagaggagc	aattggaccg	gcaggaaaag	301140
atggagcagt	tggccctcaa	ggccctccag	gagaaagagg	agaaaatgga	cgccaggaaa	301200
gagatggagc	agttggccct	caaggagaaa	gaggagcaat	tggaccggca	ggaaaggatg	301260
gagcagttgg	ccctcaaggga	gaaagaggag	caattggacc	ggcaggaaaag	gatggagcag	301320
ttggccctgc	aggccctcaa	ggagaaagag	gagaaaatgg	acgcccaggga	agagatggag	301380
cagttggccc	tgcaggccct	ccaggagaaa	gaggagcaat	tggaccggca	ggaagagatg	301440
gagcagttgg	ccctgcaggc	cctccaggag	aaagaggagc	aacaggtata	ccaggaaagg	301500
atggcgtgga	cggttctgtg	ggccctcaag	gagaaagagg	agaaaattgga	cgcccaggaa	301560
gagatggagc	agttggccct	gcaggccctc	aagggaagaag	aggagcaaca	ggacgcgcag	301620
gaaaggatgg	tgcagttggt	cctgcaggcc	ctcaaggaga	aaaaggagaa	gctggtaagg	301680
acggttctat	agggcctcaa	ggaatacaag	gcccaggagg	agagactgga	ccaccgggaa	301740
gggacggcac	tgcagcagaa	agaggagaaa	gaggcttccc	aggaccacca	ggcgaaactg	301800
gaccaccagg	aaaggatggt	gtggatgggt	ctgagggcc	tcaagggaaa	agaggagaaa	301860
caggaccogt	tggacctagg	ggtgaaccag	gtctagctgg	cctcccaggga	agagatggag	301920
caattggccc	tgcaggccct	ccaggagaaa	gaggagcaac	tggtctacca	ggaagggaatg	301980
gtgtggatgg	ttctatcggc	ccccaaggaa	gaagaggagc	aacaggccgc	gcaggaaaag	302040

WO 01/38351

84

PCT/US00/28888

```

atggggcagt tggccctgca ggcctccag gagaaagagg agcaacaggt ataccaggaa 302100
gggatggtgt ggacggttct gtgggccctc caggagaaag aggagaaact ggaccagcag 302160
gaagggacgg ttcatgtggc cctgctggcc ctcatggaga aagaggagaa aatggacgcc 302220
caggaagaga tggggcaact ggcctatag gtctgtctgg tcctcaagga gaaaaaggag 302280
aaaatggacg ccaggaaga gatggagcaa ctggccctat aggccctaga ggagaaactg 302340
gtgcaatggg aaagaatggc gtggacggtt ctatgggtcc tcaaggaaga agaggagcaa 302400
caggccgcgc aggaaggat ggggcagttg gccctgtctg ccctccagga gaaagaggag 302460
aaactggacc agcaggaagg gacggttcag ttggccctgc tggccctcaa ggagaaacag 302520
gattaactgg cagccagga agagatggag caactggccc tataggctct gctggccctc 302580
aaggagaaaa aggagaaaat ggacgcccag gaagagatgg agcaactggc cctataggtc 302640
ctgctggccc tcaaggagaa aaaggagaaa atggacgcc aggaagagat ggagcaactg 302700
gccctatagg tcctgtctgg cctcaaggag aaacaggatt aactggacgc ccaggaagag 302760
atggagcaac tggccctata ggtcctagag gagaaactgg tgcaatggga aagaatggtg 302820
tggacggttc tacgggtcct caaggaagaa gaggagcaac aggccgcgcg ggaaggatg 302880
gagcagttgg cctgtctggc cctccaggag aaagaggaga aaatggacgc ccaggaagag 302940
atggagcaac tggccctata ggtcctgtg gccctcaagg agaaacagga ttagctgggc 303000
tgccaggaag agatggagca attggtcctc aaggagaaaa gggagaaaaat ggacgcccag 303060
gaaaggatgg ggcaactggc cctatgggtc ctccaggaga aaggggagag actggtccta 303120
taggtcctgc tggccctcaa ggagcaactg gtcttcagg aagggatggt gtggatggtt 303180
ctgttggccc tcaaggaaaa agaggattaa tagggcgcac aggaagggat ggggcaattg 303240
ggcctgtagg tcctgcaggc cctaaaggag aaacaggatt agctggcctg ccaggatag 303300
atggaaagga cgggtccgtg ggtcctcaag gagcaattgg acctataggc ccacgaggag 303360
aaagaggaga aactggacga ccaggaaggg acggtgagga tggttccaca ggccctatgg 303420
gcccccaagg actaagagga gctacgggag ctccaggacc gcaaggagaa agaggattaa 303480
agggacggcc aggaaaagat ggtgaaacag gtctccagg gcgacaagga agggatggaa 303540
taatgggtcc taggggtcct cgaggagaaa aaggagcacc tggtaatgat ggtctagagg 303600
gacctgaagg aagatggtt gcacctggtc ccgctggccc tattggacct caaggaataa 303660
gaggattaaa aggtatccag ggacgaccag gaagagacgg agaaatggga ccagccggca 303720
aggacggaat agaaggccct agaggtcaag atggaacaac tggcgctaaa ggacctagag 303780
gattaagagg ttttcaagga agaacaggag aaactggtgc acaaggatct agaggagaaa 303840
aaggcgatag agggctaaca ggcctcaag gaagagacgg tccaccgggt gaagaaggte 303900
ctcaaggtct agaggagaa aggggagcac ctggccctag aggtcctaga ggtattcgtg 303960
gccgttcagg acctcaagga agtaacggcg tgcaaggacc tcgaggtccc cgaggaacaa 304020
aaggaagaac aggaatacaa ggcctcactg gcatagaagg tcctcgaggt cctagaggta 304080
tacaaggaaa ggaaggaaag atggggaaaa ttggacatcg aggagaaaag ggtgataaag 304140
gagaccgttg agaacaaggc atcgctggag cagacgggga aaaagggtcca agaggtttac 304200
gaggaattcg aggccctatt ggtgctcctg gtaagcctgg cacggaaggg gttagaggtc 304260
ctagaggggt gagaggtgtt cctggctatc ctggcgacac aggggaatta ggtccccaag 304320
gaccaacagg tcctcaaggg ccagcaggtc ctcaagggcc gatggggcgt acaggagata 304380
ctggtcccat gggccctcct ggagcagttg gaccaagagg agagaaagga ggtagaggaa 304440
gaaagggaaa aaatggccct aaaggagcgg acggaaaaga tgccgtaaat atcatacaa 304500
aatattcaat caccatgct cgtgcagaga taatgtggga aggaaatgaa atcgagaaag 304560
catacattgg aagatcttat ggaactgata caatccctgt gatgatagaa aatagaatag 304620
ggatgacaaa tgaggacaaa aaaaacgaat attgtataca agtaatgaca atgcactcaa 304680
taacaactag aggaagaaca tcgggtgttt ttgtggttaag caataagaca gattatatcc 304740
ttttagttac tttactgatg ccagaaagtg tttcctgtag aacagatgtc agtacaatg 304800
cgaggtcaga gagggtgaat gctgttagag aaagagaaaag caaatcgta agatttatta 304860
ggccgtctga ccaatctata ggtactcatt cacgttcaaa aattgccgtg gtaatgtatc 304920
cagacgcaag catgagttac tcagttgata cattagacgc tgatgtggcg cgaagagaaa 304980
caacgtctgt gcttttatta gcagaaacca tacacgggga aaaagataga ggtttctatg 305040
ctgatagagg aactgtaggg aggttgatgg tacctccac tgaagaagag ttattggtat 305100
tgcaagc                                     305107

```

<210> 2  
 <211> 855  
 <212> DNA  
 <213> SHRIMP

<400> 2  
 atggctgctg caaaaatgga cgctatcctt gcggatatta atgggaatga tacagattta 60  
 tccaagctaa tcacagacgt gattcaaaag agagccaagg ctgtcatgga tagaaatag 120



```

gctaaaatgg acatgaatag aagagtagat gaggcatttc aggaagccgt agcggccaag 180
aaacaaaaag cattagtggg atttgataaa ctcgtggaag aaactgacag cggacaaagt 240
gtccctccaa cattatcggg atccgattac gacgcgtggg tagacagagc catgccttcg 300
catattgaac ttgtagagag tgttgaggga gattctttgt atgataaaact ccctcctttt 360
aacgtacaag acatagacga ccaaatcggg gatgagatag atacaccaat atcttacctt 420
gccatggtag tggtaaaagt cgactgtgaa actggggata tcgaagaaga gtacaatctt 480
gctcctacct ttggtgtgac acaaaataat aaaatataca gagatgaaag agaccagatt 540
tttacaaagg ctgataaatac tgtgcgtatt tttaaacttg ctaaattgga tagtatatca 600
ggtaaaagta gacaactgac gtatgcggtg aaaaataaca atgaatatac agagtttgte 660
tgtagcgtct ttgcagagtt tgaatctgac agtgacacta ctaaatacagg tataggtatt 720
cgtgaatatg acaaaccctaa aaatgaattc gaatatgagg agcgagagat ctttacattt 780
tttattccaa tacaaccgcg agggacgaaa ttattgttat actttttggt ggacgtcagg 840
tcccgattta tatag                                     855

```

<210> 3  
 <211> 284  
 <212> PRT  
 <213> SHRIMP

```

<400> 3
Met Ala Ala Ala Lys Met Asp Ala Ile Leu Ala Asp Ile Asn Gly Asn
 1          5          10          15
Asp Thr Asp Leu Ser Lys Leu Ile Thr Asp Val Ile Gln Lys Arg Ala
 20          25          30
Lys Ala Val Met Asp Arg Asn Arg Ala Lys Met Asp Met Asn Arg Arg
 35          40          45
Val Asp Glu Ala Ile Gln Glu Ala Val Ala Ala Lys Lys Gln Lys Ala
 50          55          60
Leu Val Val Phe Asp Lys Leu Val Glu Glu Thr Asp Ser Gly Gln Ser
 65          70          75          80
Val Pro Pro Thr Leu Ser Gly Ser Asp Tyr Asp Ala Trp Val Asp Arg
 85          90          95
Ala Met Pro Ser His Ile Glu Leu Val Glu Ser Val Glu Gly Asp Ser
100          105          110
Leu Tyr Asp Lys Leu Pro Pro Phe Asn Val Gln Asp Ile Asp Asp Gln
115          120          125
Ile Gly Asp Glu Ile Asp Thr Pro Ile Ser Tyr Leu Ala Met Val Val
130          135          140
Val Lys Val Asp Cys Glu Thr Gly Asp Ile Glu Glu Glu Tyr Asn Leu
145          150          155          160
Ala Pro Thr Phe Gly Val Thr Gln Asn Asn Lys Ile Tyr Arg Asp Glu
165          170          175
Arg Asp Gln Ile Phe Thr Lys Ala Asp Lys Ser Val Arg Ile Phe Lys
180          185          190
Leu Ala Lys Leu Asp Ser Ile Ser Gly Lys Ser Arg Gln Leu Thr Tyr
195          200          205
Ala Val Lys Asn Asn Asn Glu Tyr Thr Glu Phe Val Cys Ser Val Phe
210          215          220
Ala Glu Phe Glu Ser Asp Ser Asp Thr Thr Lys Ser Gly Ile Gly Ile
225          230          235          240
Arg Glu Tyr Asp Lys Pro Lys Asn Glu Phe Glu Tyr Glu Glu Arg Glu
245          250          255
Ile Phe Thr Phe Phe Ile Pro Ile Gln Pro Ala Gly Thr Lys Leu Leu
260          265          270
Leu Tyr Phe Leu Val Asp Val Arg Ser Arg Ile Ile
275          280

```

<210> 4  
 <211> 1461  
 <212> DNA

<213> SHRIMP

<400> 4

```

atgttcgtca taagcatagc tacgtcattg gtgttattct tcttccttct tttcgtctca 60
ataacaattt tagatgggtgc aaaaacaatc gactctcaac cttttaggaa aaggaggaaag 120
aggaagagat atcgtaccag tgagagtggg gacggcatag acggagggaac tggaacaaca 180
aacggaggag gaggagggtg aggagaagga ggtgggtggg gaacaaatgg aaatggaaact 240
ggaacaacaa acggaggagg aggtggagga gaaggagggtg gtgggtggaac aaatggaaat 300
ggttctggaa caacaaacgg aggaggaggt ggaggagaag gaggtgggtg tggaacaaat 360
ggaggaggaa atggaaatgg aggaggaaat ggaaatggaa atggaaatgg aggggatata 420
gacacagacg attttgagcc tacgccagcc cttctaaaag aacgtctact taattccata 480
tcttcaaaac ctaaaagaata ctatgaagca ttcgtatctg cagaggtaga aactgcctta 540
caattatctc gagatgattc taccctaaact ataataattg atgacgacca attagaattg 600
gacgcctcag acactctaca aggaaaacca agggattatt tgttcaagct cgcagggtgtt 660
tcctcgccct ttttagaagg taccactatc cgcaaagcgg aagaccgtgc ccgtaataata 720
aatgaggaag aaattgcaca aacaatatata agtcaactaa gagaaaaaca catcaacgat 780
gaatacagatg gaaaaatatgc cacacccgag gaaagagctg atttttccaa tagtcttaat 840
cttgttacta aatacacaaa ccatgaagtg ggtcttttag taggagaaac tattgaaaag 900
gctttccctc atgagattga gtttgagaga tgcataattc tagtagagga ttttaataagt 960
ggaactatta cttcaaacac tatgcagtac aggtccaacg cttacaaaat cagagttagta 1020
gaaggatcaa caacagatcc aggggaagtt gtccctgatg attgtttggt ttttgccgta 1080
gtagtaaata aggaacaaca ttctctagaa atatctgcaa ctaacagatg ccaagacata 1140
tgctttgtaa ttattcctcg tttatccgca ataggaaaaa atgctaccat ggtaataagg 1200
aaaggtgatg aaattaaaca agaaacctat ctgtttgtgg ccaataagaa tgacactact 1260
catttttcaa tcatacacaga caaggatgaa tctgttgtaa tagaattaaa catgtctatt 1320
ttctcagaaa gaattctocc cactttgagt gaccctgcaa ccgttccaag acctttgact 1380
gacgccaacg tactgtcagc ctacggaaag cgcctagggtg ttgggtgcctt tacagacaaa 1440
aattttattgt ccagccaata a

```

<210> 5

<211> 482

<212> PRT

<213> SHRIMP

<400> 5

```

Met Phe Val Ile Ser Ile Ala Thr Ser Leu Val Leu Phe Phe Phe Leu
1 5 10 15
Leu Phe Val Ser Ile Thr Ile Leu Asp Gly Ala Lys Thr Ile Asp Ser
20 25 30
Gln Pro Phe Arg Lys Arg Arg Lys Arg Lys Arg Tyr Arg Thr Ser Glu
35 40 45
Ser Gly Asp Gly Ile Asp Gly Gly Thr Gly Thr Thr Asn Gly Gly Gly
50 55 60
Gly Gly Gly Gly Glu Gly Gly Gly Gly Gly Thr Asn Gly Asn Gly Thr
65 70 75 80
Gly Thr Thr Asn Gly Gly Gly Gly Gly Gly Glu Gly Gly Gly Gly
85 90 95
Thr Asn Gly Asn Gly Ser Gly Thr Thr Asn Gly Gly Gly Gly Gly
100 105 110
Glu Gly Gly Gly Gly Gly Thr Asn Gly Gly Gly Asn Gly Asn Gly Gly
115 120 125
Gly Asn Gly Asn Gly Asn Gly Asn Gly Gly Asp Thr Asp Thr Asp Asp
130 135 140
Phe Glu Pro Thr Pro Ala Leu Leu Lys Glu Arg Leu Leu Asn Ser Ile
145 150 155 160
Ser Ser Lys Pro Lys Glu Tyr Tyr Glu Ala Phe Val Ser Ala Glu Val
165 170 175
Glu Thr Ala Leu Gln Leu Ser Arg Asp Asp Ser Thr Gln Thr Ile Ile
180 185 190
Ile Asp Asp Asp Gln Leu Glu Leu Asp Ala Ser Asp Thr Leu Gln Gly
195 200 205

```

Lys Pro Arg Asp Tyr Leu Phe Lys Leu Ala Gly Val Ser Ser Ala Phe  
 210 215 220  
 Leu Glu Gly Thr Thr Ile Arg Lys Ala Glu Asp Arg Ala Arg Asn Ile  
 225 230 235 240  
 Asn Glu Glu Glu Ile Ala Gln Thr Ile Leu Ser Gln Leu Arg Glu Lys  
 245 250 255  
 His Ile Asn Asp Glu Tyr Asp Gly Lys Tyr Ala Thr Pro Glu Glu Arg  
 260 265 270  
 Ala Asp Phe Ser Asn Ser Leu Asn Leu Val Thr Lys Tyr Thr Asn His  
 275 280 285  
 Glu Val Gly Leu Leu Val Gly Glu Thr Ile Glu Lys Ala Phe Phe Ile  
 290 295 300  
 Glu Phe Glu Arg Cys Ile Ile Leu Val Glu Asp Phe Asn Ser Gly Thr  
 305 310 315 320  
 Ile Thr Ser Asn Thr Met Gln Tyr Arg Ser Asn Ala Tyr Lys Ile Arg  
 325 330 335  
 Val Val Glu Gly Ser Thr Thr Asp Pro Gly Glu Val Val Pro Asp Asp  
 340 345 350  
 Cys Leu Val Phe Ala Val Val Val Asn Lys Glu Gln His Ser Leu Glu  
 355 360 365  
 Ile Ser Ala Thr Asn Arg Cys Gln Asp Ile Cys Phe Val Ile Ile Pro  
 370 375 380  
 Arg Leu Ser Ala Ile Gly Lys Asn Ala Thr Met Val Ile Arg Lys Gly  
 385 390 395 400  
 Asp Glu Ile Lys Gln Glu Thr Tyr Leu Phe Val Ala Asn Lys Asn Asp  
 405 410 415  
 Thr Thr His Phe Ser Ile Ile Thr Asp Lys Asp Glu Ser Val Gly Ile  
 420 425 430  
 Glu Leu Asn Met Leu Ile Phe Ser Ile Leu Pro Thr Leu Ser Asp Pro  
 435 440 445  
 Ala Thr Val Pro Arg Pro Leu Thr Asp Ala Asn Val Leu Ser Ala Tyr  
 450 455 460  
 Gly Lys Arg Leu Gly Val Gly Ala Phe Thr Asp Lys Asn Leu Leu Ser  
 465 470 475 480  
 Ser Gln

<210> 6  
 <211> 903  
 <212> DNA  
 <213> SHRIMP

<400> 6  
 atgggagata agcaaaaggt ggaacagctt ttaagggaat taaaggcaga ggctaattgat 60  
 gattggctat cggctaattgt tgatccaatt gttgagagat ttgtcacaac taaatctgat 120  
 gagacagccc aagttgttaa acaggccggt gatgaaaaat acgatgaatt attagaagat 180  
 aagggtgaag aaatgagacc ggatataatc aatgaagcat ccgaaacata tgataaactt 240  
 gctgctgata tgataagaga ggtagacact agtagtggtta ttgctcctgc aatagctggc 300  
 acagtggcaa gaactatcaa taatttaaga gataaaagga aagaatatga aaagaggcta 360  
 tggacattag cctacaaacc atggagaaga tatgtacaag caattacagt gatggaattt 420  
 cgtttatcat ataaaagacct gactgtccat gccattccg atacatat t aacattccct 480  
 tttttaagaa taataaaagat cgcttatatt aataacgacc gcgccagccc cgtgaactgc 540  
 tccttgctcag tgtcttatcc caacaagtcc gagtggggaa atgataacgg agttggacgt 600  
 aaggttgaca tacacatcag aagaaatgat ttacaggaaa aggatcttta cctgtctgtt 660  
 atttgtatgt tagatactga tttttcagga tatgataaag ctgtggaagt tgatgcccat 720  
 aaatttcatt ttgaagcagg gaacaggact atgtttttac ccaaaacctc taatctattt 780  
 aatcgatcgc atatagtaaa ctctaaaata tgcactatag tgttccctcc cgcttctgcc 840  
 agttctgctt ctacaacaga actggataat gtctattaca ggatcacatg cacatgctcc 900  
 tag 903

<210> 7  
 <211> 298  
 <212> PRT  
 <213> SHRIMP

<400> 7

Met	Gly	Asp	Lys	Gln	Lys	Val	Glu	Gln	Leu	Leu	Arg	Glu	Leu	Lys	Ala
1				5					10					15	
Glu	Ala	Asn	Asp	Asp	Trp	Leu	Ser	Ala	Asn	Val	Asp	Pro	Ile	Val	Glu
			20					25					30		
Arg	Phe	Val	Thr	Thr	Lys	Ser	Asp	Glu	Thr	Ala	Gln	Val	Val	Lys	Gln
		35					40					45			
Ala	Val	Asp	Glu	Lys	Tyr	Asp	Glu	Leu	Leu	Glu	Asp	Lys	Val	Glu	Glu
	50					55					60				
Met	Arg	Pro	Asp	Ile	Ile	Asn	Glu	Ala	Ser	Glu	Thr	Tyr	Asp	Lys	Leu
65					70					75					80
Ala	Ala	Asp	Met	Ile	Arg	Glu	Val	Asp	Thr	Ser	Ser	Val	Ile	Ala	Pro
			85						90					95	
Ala	Ile	Ala	Gly	Thr	Val	Ala	Arg	Thr	Ile	Asn	Asn	Leu	Arg	Asp	Lys
			100					105					110		
Arg	Lys	Glu	Tyr	Glu	Lys	Arg	Leu	Trp	Thr	Leu	Ala	Tyr	Lys	Pro	Trp
		115					120					125			
Arg	Arg	Tyr	Val	Gln	Ala	Ile	Thr	Val	Met	Glu	Phe	Arg	Leu	Ser	Tyr
	130					135					140				
Lys	Asp	Leu	Thr	Val	His	Ala	Asn	Ser	Asp	Thr	Tyr	Leu	Thr	Phe	Pro
145					150					155					160
Phe	Leu	Arg	Ile	Lys	Lys	Ile	Ala	Tyr	Ile	Asn	Asn	Asp	Arg	Asp	Val
				165					170					175	
Asn	Cys	Ser	Leu	Ser	Val	Ser	Tyr	Pro	Asn	Lys	Ser	Glu	Trp	Gly	Asn
			180					185					190		
Asp	Asn	Gly	Val	Gly	Arg	Lys	Val	Asp	Ile	His	Ile	Arg	Arg	Asn	Asp
		195					200					205			
Leu	Gln	Glu	Lys	Asp	Leu	Tyr	Leu	Ser	Val	Ile	Cys	Met	Leu	Asp	Thr
	210					215					220				
Asp	Phe	Ser	Gly	Tyr	Asp	Lys	Ala	Val	Glu	Val	Asp	Ala	His	Lys	Phe
225					230					235					240
His	Phe	Glu	Ala	Gly	Asn	Arg	Thr	Met	Phe	Leu	Pro	Lys	Thr	Ser	Asn
				245					250					255	
Leu	Phe	Asn	Arg	Ser	His	Ile	Val	Asn	Ser	Lys	Ile	Cys	Thr	Ile	Val
		260						265					270		
Phe	Pro	Pro	Ala	Ser	Ala	Ser	Ser	Ala	Ser	Thr	Thr	Glu	Leu	Asp	Asn
		275					280					285			
Val	Tyr	Tyr	Arg	Ile	Thr	Cys	Thr	Cys	Ser						
	290					295									

<210> 8  
 <211> 2376  
 <212> DNA  
 <213> SHRIMP

<400> 8

atggaggaag	aaagtcaacg	agtacagagg	aggatcggag	tcctcccaga	agaggctgcc	60
tcccagatcc	tgaaagacac	caagttaaga	gtatcatatc	tgggcgtggg	acattggggg	120
tactctgttt	cagttataaa	aagtgtcttc	cagaaaggat	gcaggaggaa	cgatgaggat	180
ataactgcgt	ggcfaatcag	ggaggcttat	ctctattacc	atcttggatt	gaattacatt	240
gaaaacgtaa	aacctgctgc	aaaatcatta	aatacaata	tggtaaacag	aattaaaatc	300
atagctgtgg	aggataccag	tcctagaagt	atgggtggctt	ctaagtgtg	tgtgagaact	360
ctagaaaaat	atgaaaagg	gaattttagg	caaccagtt	acttgatgga	tgctgccatg	420
aggctagttc	acgcctccag	ttctagagt	tgctcccaca	tgagagcttt	gtgttgcaag	480
gaagaggaca	gtgataaatt	aggggtatt	tattatgcta	attttaatga	acttgaaact	540

<400>	9															
Met	Glu	Glu	Glu	Ser	Gln	Arg	Val	Gln	Arg	Arg	Ile	Gly	Val	Leu	Pro	
1				5					10					15		
Glu	Glu	Ala	Ala	Ser	Gln	Ile	Leu	Lys	Asp	Thr	Lys	Leu	Arg	Val	Ser	
			20					25					30			
Tyr	Leu	Gly	Val	Gly	His	Trp	Gly	Tyr	Ser	Val	Ser	Val	Ile	Lys	Ser	
		35					40					45				
Ala	Leu	Gln	Lys	Gly	Cys	Arg	Arg	Asn	Asp	Glu	Asp	Ile	Thr	Ala	Trp	
	50					55					60					
Ser	Ile	Arg	Glu	Ala	Tyr	Leu	Tyr	Tyr	His	Leu	Gln	Tyr	Ile	Glu	Asn	
65					70					75					80	
Val	Lys	Pro	Ala	Ala	Lys	Ser	Leu	Asn	Thr	Asn	Met	Val	Asn	Arg	Ile	
				85					90					95		
Lys	Ile	Ile	Ala	Val	Glu	Asp	Thr	Ser	Pro	Arg	Ser	Met	Val	Asn	Glu	
			100					105					110			
Cys	Val	Arg	Thr	Leu	Glu	Lys	Tyr	Glu	Lys	Gly	Asn	Phe	Arg	Gln	Pro	
		115					120					125				
Ser	Tyr	Leu	Met	Asp	Ala	Ala	Met	Arg	Leu	Val	His	Ala	Ser	Ser	Ser	
	130					135					140					
Arg	Val	Cys	Ser	His	Met	Arg	Ala	Leu	Cys	Cys	Lys	Glu	Glu	Asp	Ser	
145					150					155				160		
Asp	Lys	Leu	Gly	Gly	Ile	Tyr	Tyr	Ala	Asn	Phe	Asn	Glu	Leu	Glu	Thr	
				165					170					175		
Gln	Cys	Val	Ser	Ala	Val	Asn	Phe	Ser	Pro	Ile	Glu	Arg	Ile	Lys	His	

			180						185						190		
Val	Phe	Arg	Glu	Ile	Glu	Ser	Val	Lys	Leu	Gly	Lys	Lys	Ser	Val	Gln		
		195					200					205					
Leu	Leu	Asn	Leu	Arg	Ser	Val	Ala	Ala	Tyr	His	Val	Leu	Arg	Tyr	Tyr		
		210				215					220						
Gly	Asp	Lys	Val	Lys	Asp	Thr	Asn	Lys	Lys	His	Ser	Gly	Pro	Phe	Lys		
225					230					235					240		
Arg	Lys	Glu	Phe	Glu	Gln	Phe	Trp	Gly	Leu	Cys	Phe	Lys	Phe	Val	Thr		
				245				250					255				
Gln	His	Val	Lys	Thr	Asp	Pro	Glu	Leu	Arg	Cys	Tyr	Phe	Asn	Glu	Leu		
			260				265					270					
Thr	Tyr	Ala	Ile	Asn	Trp	Arg	Arg	Asp	Phe	Phe	Cys	Ser	Lys	Gly	Phe		
		275				280					285						
Phe	Arg	Glu	Glu	Ser	Leu	Phe	Leu	Thr	Ser	Ile	Val	Glu	Leu	Ile	Ile		
		290				295				300							
Ala	Met	Cys	Ile	Gly	Asp	Arg	Lys	Gln	Phe	Ala	Lys	Ile	Gln	Lys	Arg		
305					310				315						320		
Asp	Leu	Lys	Arg	Phe	Asn	Lys	Gly	Glu	Glu	Gly	Arg	Lys	Glu	Glu	Ala		
				325				330					335				
Ala	Thr	Phe	Asp	Trp	Ile	Glu	Gly	His	Val	Lys	Arg	Met	Pro	Gln	Met		
			340				345					350					
Pro	Val	Trp	Val	Leu	Asp	Lys	His	Thr	Asn	Lys	Asn	Thr	His	Gly	Val		
		355				360				365							
Ser	Phe	Ala	Leu	Glu	Ser	Ser	Met	Val	Ser	Gly	Gly	Asp	Lys	Arg	Trp		
		370				375				380							
Ser	Pro	Gly	Val	Trp	Leu	His	Ser	Tyr	Thr	Lys	Met	Arg	Leu	Asp	Ser		
385				390				395							400		
Pro	Pro	Pro	Pro	Glu	Val	Gly	Gln	Phe	Leu	Asp	Gln	Ala	Phe	Asn	Thr		
				405				410					415				
Leu	Lys	Arg	Glu	Ala	Ala	Ser	His	Cys	Val	Thr	Arg	Asn	Ile	Cys	Thr		
			420					425				430					
Thr	Thr	Gly	Phe	Ile	Lys	Ala	Ser	Ser	Phe	Thr	Ala	Asn	Ile	Asn	Ser		
		435				440				445							
Glu	Pro	Met	Glu	Ile	Lys	Glu	Glu	Ile	Lys	Lys	Arg	Lys	Ile	Glu	Ile		
		450				455				460							
Lys	Asp	Asp	Asn	Thr	Thr	Ala	Thr	Val	Thr	Val	Ser	Ala	Thr	Thr	Ser		
465				470				475							480		
Ser	Ser	Ile	Thr	Ser	Thr	Pro	Pro	Pro	Thr	Lys	Lys	Gln	Lys	Thr	Thr		
				485				490					495				
Pro	Ser	Gly	Ser	Asn	Lys	Val	Asp	Ser	Ile	Gln	Leu	Asn	Asn	Leu	Pro		
			500				505					510					
Thr	Leu	Asn	Met	Glu	Asp	Leu	Asp	Arg	Val	Leu	Glu	Val	His	Asn	Gln		
		515				520				525				</			

WO 01/38351

91

PCT/US00/28888

Asn Thr Ser Asn Ile Leu Val Gly Asp Gly Gly Asn Leu Phe Ser Val  
675 680 685  
Asp Glu Asn Tyr Val Gly Ala Lys Asp Pro Arg Thr Ala Leu Glu Asn  
690 695 700  
Arg Lys Ile Lys Glu Leu Gln Leu Leu Leu Lys Thr Ser Phe Lys Val  
705 710 715 720  
Asn Lys Val Thr Lys Glu Asp Ile Asp Ser Cys Leu Pro Ser Trp Leu  
725 730 735  
Phe Asp Thr Ser Lys Ser Asp Lys Ile Met Asn Gly Val Cys Asn Ile  
740 745 750  
Gly Lys Asn Met Gly Ile Gly Pro Thr Thr Leu Asp Ile Val Lys Asn  
755 760 765  
Asn Cys Thr Cys Ile Leu Gly Val Val Asn Asp Leu Leu Tyr Asp Asn  
770 775 780  
Lys  
785

<210> 10  
<211> 1857  
<212> DNA  
<213> SHRIMP

<400> 10  
atggcttttaa aagacgcatt caccgagcgg ttagttgtga acaaagttgg ctccggtaca 60  
gacatggctc cagttgttga agatgaccgc cagaaatctc tcttccaaaa agtggaataa 120  
ctttatagag ttctggtagt tgaacagaaa aattctgcaa ttactctctc tggaaacaaa 180  
aacacaaaaca aaagacaatg ccgccaggta gaagaagaca aagtgatttt tgaaggagag 240  
gacagaactg tatctaattt accccaagct gttaaagaaa ccatcgcagc caatgcagaa 300  
tctatactcg attactggta caagaacggt atacctcttc ttgacacgaa gaaagaaaga 360  
agtggggaaaa gtgatacggt tctgaggact gcagtcactc gtcttgtgag atgttgtgtg 420  
tcatacaaaag acatgaaaac ttgttcgctg atatatgaat tcgaacacaa aattttaaat 480  
aagagtacgt tggacccttt attgaaggac attctcgaca acaacaaga actactacat 540  
atggattcga aatatgggag taaaactaca tccctgagc tggcaaagga aaccatcgag 600  
gccttgtaca cgactgtata taatcactgg acaaacgcct tcaagctcta ccaagcctcg 660  
ctgactcata acctgtttac gggtaaagaa tatgcatcag tcatacactt tataaggaca 720  
tggaggaaga tcgtcaaggc ttacgtgtcc aaacacaaca acgtggagag ggatctttca 780  
ctgaagaaca taatgaagaa tgagagtgtc gataatgcta atgttttaac tattgaaaaa 840  
atgtacaaga aaattggtaa cagtgtcaag aacactaata ataatagtgc ccaccaaag 900  
tcggacagtg aagatgatga tgatgatgat gacgatgact gtgaaggcat ggatgtttgt 960  
gatgaagcat cagaaaggga aaagaagcac caagaatctc tatacccaat caacactcca 1020  
gtcacaacaa tcactggcga ttacatcttt aaagtgtctc tagagcttgt tctttcacc 1080  
cacatccacc ctgaatggaa aatacctatg tgtgattttg ttaataggaa tatcccaaaa 1140  
ttaatgaagg ccatggaaac tgacatatca aatgctgtaa tagaagttag agcttcaaag 1200  
gtgaaccctg tacagatttt acccatcgct gctaacttct gggatttttg taagagcggg 1260  
aaacctcctt ctgatgtcaa attttgtatg atgtttaatg aaccatcttc taacgagacc 1320  
ctttcttctg gtgcgggtgt gtttggtcga ttattggag gaccttttc ccacaagag 1380  
aaagaatttg acattatttc aaactgtcta agaagtttat tgttgaataa agaagcggac 1440  
aatctatcta ccagaatttg gagagaagga ggttctgttg tctgttttaa ttactgcccc 1500  
attactgccg gaggtgtgt tttagggtat ggtgaacaat tgtccgagag atctattaaa 1560  
gctctgtggg ccaagaaaat ccaggacgct gtaacagaat ctgtcaagag gcaaaggaa 1620  
gcggctgaca aaaaagttag gaactgtgat ttgttaggag acgaggggtg agtatccatg 1680  
aaaactgtaa cgtttggatg tgcaaatatg ttgaaaaccc agaacggtat ggggaaattt 1740  
aatgtagttg taagcttcga ggattctatt caggctaata aggaaggggc ggccagacag 1800  
tatatgtctc agcaagtttt tactcactca tttccagctc tagatcaagg aaaataa 1857

<210> 11  
<211> 610  
<212> PRT  
<213> SHRIMP

&lt;400&gt; 11

```

Met Ala Leu Lys Asp Ala Phe Thr Glu Arg Leu Val Val Asn Lys Val
 1      5      10      15
Gly Ser Gly Thr Asp Met Ala Pro Val Glu Asp Asp Arg Gln Lys
 20      25      30
Ser Leu Phe Gln Lys Val Glu Asn Leu Tyr Arg Val Leu Val Val Glu
 35      40      45
Gln Lys Asn Ser Ala Ile Thr Leu Ser Gly Asn Lys Asn Thr Asn Lys
 50      55      60
Arg Gln Cys Arg Gln Val Glu Glu Asp Lys Val Ile Phe Glu Gly Glu
 65      70      75      80
Asp Arg Thr Val Ser Asn Leu Pro Gln Ala Val Lys Glu Thr Ile Ala
 85      90      95
Ala Asn Ala Glu Ser Ile Leu Asp Tyr Trp Tyr Lys Asn Val Ile Pro
 100     105     110
Leu Leu Asp Thr Lys Lys Glu Arg Ser Gly Lys Ser Asp Thr Phe Leu
 115     120     125
Arg Thr Ala Val Ile Cys Leu Val Arg Cys Cys Val Ser Tyr Lys Asp
 130     135     140
Met Lys Thr Cys Ser Leu Ile Tyr Glu Phe Glu His Lys Ile Leu Asn
 145     150     155     160
Lys Ser Thr Leu Asp Pro Leu Leu Lys Asp Ile Leu Asp Asn Lys Gln
 165     170     175
Glu Leu Leu His Met Asp Ser Lys Tyr Gly Ser Lys Thr Thr Ser Pro
 180     185     190
Glu Leu Ala Lys Glu Thr Ile Glu Ala Leu Tyr Thr Thr Val Tyr Asn
 195     200     205
His Trp Thr Asn Ala Phe Lys Leu Tyr Gln Ala Ser Leu Thr His Lys
 210     215     220
Pro Val Thr Gly Lys Lys Tyr Ala Ser Val Ile His Phe Ile Arg Thr
 225     230     235     240
Trp Arg Lys Ile Val Lys Ala Tyr Val Ser Lys His Asn Asn Val Glu
 245     250     255
Arg Asp Leu Ser Leu Lys Asn Ile Met Lys Asn Glu Ser Ala Asp Asn
 260     265     270
Ala Asn Val Leu Thr Ile Glu Lys Met Tyr Lys Lys Ile Gly Asn Ser
 275     280     285
Val Lys Asn Thr Asn Asn Asn Ser Ala His Gln Met Ser Asp Ser Glu
 290     295     300
Asp Asp Asp Asp Asp Asp Asp Asp Cys Glu Gly Met Asp Val Cys
 305     310     315     320
Asp Glu Ala Ser Glu Lys Lys His Gln Glu Ser Leu Tyr Pro Ile Asn
 325     330     335
Thr Pro Val Thr Thr Ile Thr Gly Asp Tyr Ile Phe Lys Val Leu Leu
 340     345     350
Glu Leu Val Leu Ser Pro His Ile His Pro Glu Trp Lys Ile Pro Met
 355     360     365
Cys Asp Phe Val Asn Arg Asn Ile Pro Lys Leu Met Lys Ala Met Asp
 370     375     380
Ile Ser Asn Ala Val Ile Glu Val Arg Ala Ser Lys Val Asn Pro Val
 385     390     395     400
Gln Ile Leu Pro Ile Ala Ala Asn Phe Trp Asp Phe Cys Lys Ser Gly
 405     410     415
Lys Pro Pro Ser Asp Val Lys Phe Cys Met Met Phe Asn Glu Pro Ser
 420     425     430
Ser Asn Glu Thr Leu Ser Ser Gly Ala Gly Val Phe Gly Arg Phe Ile
 435     440     445
Gly Gly Pro Phe Ser His Lys Ser Lys Glu Leu Asp Ile Ile Ser Asn
 450     455     460
Cys Leu Arg Ser Leu Leu Leu Asn Lys Glu Ala Asp Asn Leu Ser Thr
 465     470     475     480

```



Arg Ile Trp Arg Glu Gly Gly Ser Val Val Cys Phe Asn Tyr Cys Pro  
 485 490 495  
 Ile Thr Arg Ala Val Leu Gly Tyr Gly Glu Gln Leu Ser Ser Ile Lys  
 500 505 510  
 Ala Leu Trp Ala Lys Lys Ile Gln Asp Ala Val Thr Glu Ser Val Lys  
 515 520 525  
 Arg Gln Arg Asn Ala Ala Asp Lys Asn Ser Arg Asn Cys Asp Leu Leu  
 530 535 540  
 Gly Asp Glu Gly Val Val Ser Met Lys Thr Val Thr Phe Gly Cys Ala  
 545 550 555 560  
 Asn Met Leu Lys Thr Gln Asn Gly Met Gly Lys Phe Asn Val Val Val  
 565 570 575  
 Ser Phe Glu Asp Ser Ile Gln Ala Asn Lys Glu Gly Ala Ala Arg Gln  
 580 585 590  
 Tyr Met Ser Gln Gln Val Phe Thr His Ser Phe Pro Ala Leu Asp Gln  
 595 600 605  
 Gly Lys  
 610

<210> 12  
 <211> 5811  
 <212> DNA  
 <213> SHRIMP

<400> 12  
 atggggagta aacgtccgtg ctccagcggc caagagccag ttaccaaaaa acaaaagaag 60  
 aacaacaaca ataattcaaa tccagttcca gttatcaaca tcaaatcata cccctttctg 120  
 gctaccagaa cccaggtgct cagaagtgc gttgcagctg cagcggcatc cccttctggg 180  
 tcttcttctt cctcctcctc ctctgcttct gcggtcaagc tcccagatac gtgcaaggag 240  
 gccagaaagg tattgagcac tgtttccttg caacagtctc tggcagtgc gtatctgtgc 300  
 aactcgattt ctgtcagtta tgctggagga ggaatatctg tgttccatct ggttggtctt 360  
 cctggagctg ggaaaactac aatggtcaag gaactaattg cagttcttaa cgaccacggc 420  
 ttgattgatt ctggatctgc agacatgctc ctgtgttgca aatctaattc tgcaaaggag 480  
 agtctcatgt gtgcatgcaa gaaaccagga ggttcttctt tgatgtacct agaaagtgtg 540  
 ttttctaccc tcaacaaagg atttgagatt ccagttatat ttaggaaaga tgaataaact 600  
 ctagaaaaaa tccaatttgt ggccgataaa ctaaaatgga aagtgatata agtgttgga 660  
 aatttaaggt tccttggtat cgatgagtat acaatggcca gttgcaggga attggtgttt 720  
 attgatgctg ttttacgtat agcgaacac aggccagata tacctttcgg aggagtgttt 780  
 gttatactac ttggagataa caggcaaaat tctgcagtag tggaagataa cactaaccat 840  
 atccagaaga aaataaagaa cccgagtgc gaagaaaaac ctcaagaaga caacaagaac 900  
 aacaagaaca agaagaagaa gaaagagaag gaggaagaag aagaagaaag tgacgatgaa 960  
 gacgaaaacg aagaagaaga gggggaggag accttcttcc aagggtcagt cgaacaagat 1020  
 gctgagacga agaaagagga agaaaagagt accttcttcc aagggtcagt cgaacaagat 1080  
 aattttggcc aagaagataa tgcaaaaacta tacactgaag tttttataaa gatactgaaa 1140  
 atgttttgct ctgagagact ttttggaac ccaagtaacc tgcggaatat tgtcaataaa 1200  
 cgccatgaag caattttaat gaagagtaat aacgtcaagt cagtttaaca caatttagtc 1260  
 tcaagtgcata ttaaggttga agactgtggt aatgcttcta ataaaaagga agtgacggca 1320  
 ccaagttctt ccccggccca atccacggca gaagaaaatt gtgacgaatt tgatgacgaa 1380  
 gaagatgacg atttctttta caatgaagca tttttgaaac tgatggaaag aaatgcactt 1440  
 gaaaaggaca gagcatcggg agctttgaat ggattctctt tacgctgcaa gagcatttct 1500  
 gatgcaaacg aaaagataag aagtgaaca acttctgtat cagacaagaa gtcttctttg 1560  
 gatatgatga aatctctccc cttttccgct ctaatagaag aaggaatatg ttcagagcta 1620  
 gctcacatat ctgaactcaa gaaaatgagc aatgcaaatc tagaaaagta cactgaaaat 1680  
 gtatgcagta ttgtgttcga catgatggcc aaagccatga gagaaattga ttacagtggc 1740  
 cgagaaaaat tgtatattgt ttctagccta tcagagagat ttaaggatac acatctgacc 1800  
 tccttaatgg atgaagagat attgaacgtc aagtatgtgc acggttctga tcccaagtgt 1860  
 atagatgcag taccattcaa ttctgcacat aaccgtgcgt ctgctgtggc agcgtgtgtg 1920  
 aggaacgcct tcttttagaga tggaaaagac tttgtggatg aaactcctat cgctaattac 1980  
 tttaaagata atttgcttac agttgcttca ttttttagaga atgaaacttt aacatacaaa 2040  
 gaactactag cgaaatctga aaatataaga tctattctct tgaagaaaga aactggtaat 2100

aataatagtg	catcatcacg	tacagctgca	gcagcagctt	attatgagga	cgacgaggac	2160
tattgttatt	tcgatgaaga	agaagcaatg	gatttgggaag	acggaggaag	tggaggttct	2220
ggtatgaaaa	gtagtgggtg	gggaggtgat	gacgatgatg	aagaaagtgg	agaaatgata	2280
tatagaactg	atataccaga	taaattgcac	cgtgacgcat	ctactctaga	tcgtgtaggt	2340
catttagttg	actttcatgt	tgtgtggaaa	aaatggctca	ctgaaaataa	gccttcagat	2400
ttggtacgag	cccgtgtgtg	gtatttgtat	acactagtga	gaatgcaaca	ggttaaattt	2460
gataatggaa	aactcccttc	tcttgataaa	tctgactgtg	ctggacgctt	gttccactcg	2520
ccctcagaat	gggccacaa	cacaggagta	ggagtaggag	gaggaggagg	agcaggggca	2580
gataagcctc	tccatgacga	atattggttg	agggtattat	ctatgcccac	atcaacagga	2640
ggagatgtag	ggaagagtat	gttattacct	gcataattctt	cctatctttc	tgccctttcg	2700
agaacatata	tcagtgtcgtc	gttaaagaga	atagacatta	ttaagcacgc	ttattcgtta	2760
atgtacggta	tttcattggt	tgatatgact	gccaaccttc	aagacttggt	agatacgcga	2820
atggccggta	gatcatccag	aaatgggagt	gttttcatgg	acaactttga	cccagtgcga	2880
tattttgata	acatcttccc	ttcaatggtt	aacgaatttc	ttatgtacag	aaaggaagat	2940
gttttcaata	atggtcaaat	gatggagggt	gttaaagggt	gtttaaagat	ttccagagtg	3000
cttcaaacgt	ctcacactga	aaataacaac	attaacaata	ggcacaataa	ctcacttaaa	3060
tactctgaaa	aatctatagt	gttggccatg	caaatggtga	ctagtatttc	aaagggaaat	3120
gagaggagga	agaagattga	agagtttata	acaaaagaac	aagggcagcc	taaggatatg	3180
tgtgaacgcc	tcattggccaa	ctctaaagcc	aaacaagaaa	aggatgcaat	ttcttccaaa	3240
actgacaaga	tgatgggtgc	aatcacactc	actaaaaaac	acgtttttaa	gaatgcagtc	3300
tcgaattttg	tggatacttc	tatcatcaaa	gaaactaaaa	agaacaataa	caattcctca	3360
tcctcctctt	ctacttcact	tgcagcggct	gcggcagttg	aaaattctgt	gcccgcatta	3420
agagtggaa	taaaatttgt	agttttgaat	atggatctgt	ctgacattag	ccacgagaaa	3480
accatctctc	acaagtaccg	gcaacaatta	attaacgcca	ttaaaactcg	cagtacacca	3540
ctttttgaca	agttttacaga	caggaaaatt	ttgagggcag	ccgaatctcc	tagggcatta	3600
actacgattc	ttttggatga	gaaaaagaag	gtaactaggg	ctaaatctat	tactctttat	3660
cagggccaaa	atgtgtattt	cactacatct	aatagaatga	tccatggcac	tcaggaaagg	3720
ttcgtgacta	aggacactgg	agtgtttact	aatttaattgt	acaagaatgg	tgaattgaca	3780
gtatttgtgt	acgtggaacg	gctaggacaa	aatgccttag	aaattaaaga	aggacgacaa	3840
attatcggtg	atccaaatat	caagaatgga	ggatttggga	acaatgtata	cgttcaatat	3900
cttccatttg	aatctagtca	agccatgacg	atatattctt	gccaagggca	cactttcttt	3960
agagacacta	ttgttgacct	ttctggtgca	tctaccagg	atgctatgt	tgctgttaac	4020
cgcaatagta	atcctcaaaa	cttgttcatt	attcaaaatc	attctgttga	acggggtaac	4080
ttgtgttaaca	ttaaattgogc	catgagtaaa	gacaaggctt	atactatgcc	tattggaggc	4140
attgctgatt	ttaatgggag	cgactttatt	aaccacgata	cagttagtgt	ctctagagag	4200
tgggcagaat	cttccgctgc	aatggatgat	gactataatg	gagatggagg	agtaacaatg	4260
tattctgcct	atgacccttc	taaagatgtg	gtagctgccc	ctgaagaatt	catcctaagt	4320
agatcgggaa	aatctttatc	gtttaatgag	tcttggtagg	ccaacactgc	taaggttatt	4380
cagcaacatg	ggctagaaac	agaacttaaa	aacattagag	actttttctt	tgggtgtta	4440
aatggagatg	tggctaaaca	ttatgagaaa	ttatgtaaca	agaaaatgat	tgagctatac	4500
actgctatcg	ttcgttcaat	cacacactac	tctatagcca	gtggtattgt	aaaacagccc	4560
tcttcaaagt	tgtgtgaaga	atatgagacg	aaacaaaaga	acaagaaaaga	ctatatcaaa	4620
atacaccag	tttttgtgaa	tagggcacca	aaggagtcca	ctattgaaat	gttgttgttt	4680
gatattgcac	cccataacaa	ggctacaatt	gtgttccaat	tttatgtcca	ttatatattt	4740
ttagtgtatg	aaaaattgaa	cgttttgaat	tcttcattcg	cctttttgcc	atctcccaat	4800
ccgtgcctga	atcagtatgt	acgccccaaa	tcaattacca	ccaattctac	tcagtaccac	4860
aatctgggtt	acgagtcctaa	agatttttgc	cattgcaagg	acgggggaga	aagagatgtc	4920
aaattgcgtc	ttcctataac	aagcgcggat	gagttttcaa	ataacattga	aggtattttg	4980
aaaaaagtta	gtgacacatc	taatcaaaat	aaggtttaata	aatacatgga	tggtgtatgt	5040
aagagtatgc	agcacaactt	gcgtaggact	ggtaaatttt	gtcgtcctac	ggaaacttgt	5100
ggtctatcga	aacatgggtc	tattgtaacg	tccacatgca	cagctcaaga	gaaggagaa	5160
aatatccatg	tagatgcaga	aaagggatgg	ctgtgtatgt	cagacgaagc	taatgtatat	5220
tgtatgttaa	tgttcatgtc	aaaaatagca	gcagcttctg	gtgttagtga	attccccatt	5280
aaggacaaaa	gtattcttga	aagtaaccca	gaaacaccat	ctgatacgat	tagtcttttg	5340
gcaccaagaa	agacaatttc	acctactaac	aatctccact	tttcaatgtc	tgaagatgtc	5400
ctgttctgtg	gacaagttca	tccaatgaaa	agggtaaca	tttcttttgc	tgtaaagaga	5460
actggaggag	cactgaaaag	tacttttgag	gaagaagaag	gtcttcccac	aaagatatct	5520
tctcccaact	ttgccacata	tcctctattc	aaaaaatgta	aaatgtatgg	tgctataatc	5580
attgccatga	ctgaaatgca	agggcatgaa	tttgccaagt	attcaacatt	ggatatcaga	5640
aagagcatgt	ttactggtgt	tggaaactgtg	gtggatttgg	aaaagatatc	aggagaaggt	5700
aatgaagtaa	tggataaagt	tgacaaaattt	attgtgaaaa	atgtatcaaa	catattgttt	5760

aaggaacaag ggaagcgtgt ctctttcttt gtttcatgtg ctatacacta g

5811

<210> 13

<211> 1926

<212> PRT

<213> SHRIMP

<400> 13

Met	Gly	Ser	Lys	Arg	Pro	Cys	Ser	Ser	Gly	Gln	Glu	Pro	Val	Thr	Lys
1				5					10					15	
Lys	Gln	Lys	Lys	Asn	Asn	Asn	Asn	Asn	Ser	Asn	Pro	Val	Pro	Val	Ile
			20					25					30		
Asn	Ile	Lys	Ser	Tyr	Pro	Phe	Leu	Ala	Thr	Arg	Thr	Gln	Val	Leu	Arg
		35					40					45			
Ser	Ala	Val	Ala	Ala	Ala	Ala	Asp	Ser	Gly	Ser	Ser	Ser	Ser	Ser	Ser
		50					55				60				
Ser	Ser	Ala	Ser	Ala	Val	Lys	Leu	Pro	Asp	Thr	Cys	Lys	Glu	Ala	Arg
65					70					75				80	
Lys	Val	Leu	Ser	Thr	Val	Ser	Leu	Gln	Gln	Ser	Leu	Ala	Val	Arg	Tyr
				85					90					95	
Leu	Cys	Asn	Ser	Ile	Ser	Val	Ser	Tyr	Ala	Gly	Gly	Gly	Ile	Ser	Val
			100					105					110		
Phe	His	Leu	Gly	Gly	Leu	Pro	Gly	Ala	Gly	Lys	Thr	Thr	Met	Val	Lys
		115					120					125			
Glu	Leu	Ile	Ala	Val	Leu	Asn	Asp	His	Gly	Leu	Ile	Asp	Ser	Gly	Ser
		130				135					140				
Ala	Asp	Met	Leu	Leu	Cys	Cys	Lys	Ser	Asn	Ser	Ala	Lys	Glu	Ser	Leu
145					150					155					160
Met	Cys	Ala	Cys	Lys	Lys	Pro	Gly	Gly	Ser	Ser	Leu	Met	Tyr	Pro	Glu
				165					170					175	
Ser	Val	Phe	Ser	Thr	Leu	Asn	Lys	Gly	Phe	Glu	Ile	Pro	Val	Ile	Phe
			180					185					190		
Arg	Lys	Asp	Glu	Ile	Thr	Leu	Glu	Lys	Ile	Gln	Phe	Val	Ala	Asp	Lys
		195					200					205			
Leu	Lys	Trp	Lys	Val	Ile	Gln	Val	Leu	Ala	Asn	Leu	Arg	Phe	Leu	Val
		210				215					220				
Ile	Asp	Glu	Tyr	Thr	Met	Ala	Ser	Cys	Arg	Glu	Leu	Val	Phe	Ile	Asp
225					230					235					240
Ala	Val	Leu	Arg	Ile	Ala	Lys	His	Arg	Pro	Asp	Ile	Pro	Phe	Gly	Gly
				245					250					255	
Val	Phe	Val	Ile	Leu	Leu	Gly	Asp	Asn	Arg	Gln	Asn	Ser	Ala	Val	Val
			260					265					270		
Glu	Asp	Asn	Thr	Asn	His	Ile	Gln	Lys	Lys	Ile	Lys	Asn	Pro	Ser	Glu
		275					280					285			
Glu	Glu	Lys	Pro	Gln	Lys	Asn	Asn	Lys	Asn	Asn	Lys	Asn	Lys	Lys	Lys
		290				295					300				
Lys	Lys	Glu	Lys	Lys	Glu	Lys	Gly	Gly	Glu	Glu	Glu	Glu	Gly	Asp	Glu
305					310					315					320
Asn	Glu	Glu	Glu	Glu	Gly	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Ser	Asp
				325					330					335	
Asp	Glu	Ala	Glu	Thr	Lys	Lys	Glu	Glu	Glu	Lys	Ser	Thr	Phe	Phe	Gln
			340					345					350		
Gly	Ser	Val	Glu	Gln	Asp	Asn	Phe	Gly	Gln	Glu	Asp	Asn	Ala	Lys	Leu
		355					360					365			
Tyr	Thr	Glu	Val	Phe	Ile	Lys	Ile	Leu	Lys	Met	Phe	Cys	Ser	Arg	Asp
		370				375					380				
Phe	Phe	Gly	Asn	Pro	Ser	Asn	Leu	Arg	Asn	Ile	Val	Asn	Lys	Arg	His
385					390					395					400
Glu	Ala	Ile	Leu	Met	Lys	Ser	Asn	Asn	Val	Lys	Ser	Val	Asn	Asn	Asn
				405					410				415		
Leu	Val	Ser	Ser	Ala	Ile	Lys	Val	Glu	Asp	Cys	Gly	Asn	Asn	Lys	Lys





```

1395      1400      1405
Gly Val Thr Met Tyr Ser Ala Tyr Asp Pro Ser Lys Asp Val Val Ala
1410      1415      1420
Ala Ala Glu Glu Phe Ile Leu Ser Arg Ser Gly Lys Ser Leu Ser Phe
1425      1430      1435      1440
Asn Ala Ser Trp Met Ala Asn Thr Ala Lys Val Ile Gln Gln His Gly
1445      1450      1455
Leu Glu Thr Glu Leu Lys Asn Ile Arg Asp Phe Phe Phe Gly Val Asn
1460      1465      1470
Asn Gly Asp Val Ala Lys His Tyr Glu Lys Leu Cys Asn Lys Lys Met
1475      1480      1485
Ile Glu Leu Tyr Thr Ala Ile Val Arg Ser Ile Thr His Tyr Ser Ile
1490      1495      1500
Ala Ser Gly Ile Val Lys Gln Pro Ser Ser Lys Leu Cys Glu Glu Tyr
1505      1510      1515      1520
Glu Thr Lys Gln Lys Asn Lys Lys Asp Tyr Ile Lys Ile His Pro Val
1525      1530      1535
Phe Val Asn Arg Ala Pro Lys Glu Ser Thr Ile Glu Met Leu Leu Phe
1540      1545      1550
Asp Ile Ala Pro His Asn Lys Ala Thr Ile Val Phe Gln Phe Tyr Val
1555      1560      1565
His Tyr Ile Phe Leu Val Tyr Glu Lys Leu Asn Val Leu Asn Ser Ser
1570      1575      1580
Phe Ala Phe Leu Pro Ser Pro Asn Pro Cys Leu Asn Gln Tyr Val Arg
1585      1590      1595      1600
Pro Lys Ser Ile Thr Thr Asn Ser Thr His Val Pro Asn Leu Gly Tyr
1605      1610      1615
Glu Ser Lys Asp Phe Ala His Cys Lys Asp Gly Gly Glu Arg Asp Val
1620      1625      1630
Lys Leu Arg Leu Pro Ile Thr Ser Ala Asp Glu Phe Ser Asn Asn Ile
1635      1640      1645
Glu Gly Ile Leu Lys Lys Val Ser Asp Thr Ser Asn Gln Asn Lys Val
1650      1655      1660
Asn Lys Tyr Met Asp Val Val Cys Lys Ser Met Gln His Asn Leu Arg
1665      1670      1675      1680
Arg Thr Gly Lys Phe Cys Arg Pro Thr Glu Thr Cys Gly Leu Ser Lys
1685      1690      1695
His Gly Ser Ile Val Thr Ser Thr Cys Thr Ala Gln Glu Lys Gly Glu
1700      1705      1710
Asn Ile His Val Asp Ala Glu Lys Gly Trp Leu Cys Met Ser Asp Glu
1715      1720      1725
Ala Asn Val Tyr Cys Met Leu Met Phe Met Ser Lys Ile Ala Ala Ala
1730      1735      1740
Ser Gly Val Ser Glu Phe Pro Ile Lys Asp Lys Ser Ile Ser Asn Pro
1745      1750      1755      1760
Glu Thr Pro Ser Asp Thr Ile Ser Leu Leu Ala Pro Arg Lys Thr Ile
1765      1770      1775
Ser Pro Thr Asn Asn Leu His Phe Ser Met Ser Glu Asp Val Leu Phe
1780      1785      1790
Cys Gly Gln Val His Pro Met Lys Arg Val Gln Phe Ser Leu His Val
1795      1800      1805
Lys Arg Thr Gly Gly Ala Leu Lys Ser Thr Phe Glu Glu Glu Gly
1810      1815      1820
Leu Pro Thr Lys Ile Phe Ser Pro Asn Phe Ala Thr Tyr Pro Leu Phe
1825      1830      1835      1840
Lys Lys Cys Lys Met Tyr Gly Ala Ile Ile Ile Ala Met Thr Glu Met
1845      1850      1855
Gln Gly His Glu Phe Ala Lys Tyr Ser Thr Leu Asp Ile Arg Lys Ser
1860      1865      1870
Met Phe Thr Gly Val Gly Thr Val Val Asp Leu Glu Lys Ile Ser Gly
1875      1880      1885

```



WO 01/38351

100

PCT/US00/28888

210	215	220
Thr Lys Asn Thr Arg Asn Ala Lys Phe Ser Leu Arg Tyr Arg Asn Glu		
225	230	235
Lys Pro Val Asp His Leu Leu Leu Tyr Cys Met Val Thr Tyr Phe		240
	245	255

<210> 16  
 <211> 348  
 <212> DNA  
 <213> SHRIMP

<400> 16  
 atggcttcgc cattggtggc gtcttttagga ggaggaaaaa atatactttt tggttttactt 60  
 ttaataacaa taattgtcat tgttattgct gtaataataa tcaaggctcc tctcttagca 120  
 tctctcctgg cgggcacagc actggccggc acgatcgctt ctgcgctggg atcaatacca 180  
 ggagtgggag gtgcattcaa gaaagccttt ggaaaaggaa aggggaaagg aggacaaaaa 240  
 acaccagatg gtggggccaa gaaaactaac caaaaaccaa agaagggtaa gaaaaaaccc 300  
 cccaccgctc gttccatatt caagaggata cctaaaataa aatttttag 348

<210> 17  
 <211> 111  
 <212> PRT  
 <213> SHRIMP

<400> 17  
 Met Asp Leu Val Ala Ser Leu Gly Gly Gly Lys Asn Ile Leu Phe Gly  
 1 5 10 15  
 Leu Leu Leu Ile Thr Ile Ile Val Ile Val Ile Ala Val Ile Ile Ile  
 20 25 30  
 Lys Ala Pro Leu Leu Ala Ser Leu Ala Gly Thr Ala Gly Thr Ile  
 35 40 45  
 Ala Ser Ala Leu Gly Ser Ile Pro Gly Val Gly Gly Ala Phe Lys Lys  
 50 55 60  
 Ala Phe Gly Lys Gly Lys Gly Lys Gly Gly Pro Lys Thr Pro Asp Gly  
 65 70 75 80  
 Gly Ala Lys Lys Thr Asn Gln Lys Pro Lys Lys Gly Lys Lys Lys Pro  
 85 90 95  
 Pro Thr Arg Arg Ser Ile Phe Lys Arg Ile Pro Lys Ile Lys Phe  
 100 105 110

<210> 18  
 <211> 687  
 <212> DNA  
 <213> SHRIMP

<400> 18  
 atggctctctt ctagaacatc aacaacatct tcatctgcag tagcagccac ctctactctt 60  
 ctccccacca agaggaagag ggagccagaa gaagtaaagg tgaaagtgga agtaaaaatg 120  
 gaacaagaag aactggtaga ggactcgtca agtaacaagc gccccagaat taaggaagag 180  
 aaggaggaag aacacaaaaga aacacatcac ctctccctcc catgtaaaga agaagaagac 240  
 gatggtgaag aagaggaata tgaggaagag gaggatgagg aagaatatga agacagagt 300  
 gacgacgaca ctgcagagaa aatggaaaat cttttggtgc aactggacaa tactacaaa 360  
 aacatcaaac tgaaaaaccc cctaagggaa catgacatgg cagtttcaca ctatgagcat 420  
 gaatttgagg taaaaaatac tgtcaatttt agttttggag tactatctga tattgggttc 480  
 ctgatcaacc gtgaagccgt ttctaggtgg ggtaatacac cccacccaaa agagtttggc 540  
 gacatggaga ttggatctct tacagttaac cagttgctcc acaagtgtga taattttgta 600  
 caggctgtag tacagaaagt gaaggaagat ataaccctt ctattgaagt tacaatagat 660  
 agtttgattg atgaccttg ttggtaa 687



WO 01/38351

101

PCT/US00/28888

<210> 19  
<211> 228  
<212> PRT  
<213> SHRIMP

<400> 19  
Met Val Ser Ser Arg Thr Ser Thr Thr Ser Ser Ser Ala Val Ala Ala  
1 5 10 15  
Thr Ser Thr Leu Leu Pro Thr Lys Arg Lys Arg Glu Pro Glu Glu Val  
20 25 30  
Lys Val Lys Val Glu Val Lys Met Glu Gln Glu Glu Leu Val Glu Asp  
35 40 45  
Ser Ser Ser Asn Lys Arg Pro Arg Ile Lys Glu Glu Lys Glu Glu Glu  
50 55 60  
His Lys Glu Thr His His Leu Ser Leu Pro Cys Lys Glu Glu Glu Asp  
65 70 75 80  
Asp Gly Glu Glu Glu Glu Tyr Glu Glu Glu Glu Asp Glu Glu Glu Tyr  
85 90 95  
Glu Asp Arg Val Asp Asp Asp Thr Ala Glu Lys Met Glu Asn Leu Leu  
100 105 110  
Val Gln Leu Asp Asn Thr Thr Lys Asn Ile Lys Leu Lys Asn Pro Leu  
115 120 125  
Arg Glu His Asp Met Ala Val Ser His Tyr Glu His Glu Phe Glu Val  
130 135 140  
Gln Asn Thr Val Asn Phe Ser Phe Gly Val Leu Ser Asp Ile Gly Phe  
145 150 155 160  
Leu Ile Asn Arg Glu Ala Val Ser Arg Trp Gly Asn Thr Pro Pro Pro  
165 170 175  
Lys Glu Phe Gly Asp Met Glu Ile Gly Ser Leu Thr Val Asn Gln Leu  
180 185 190  
Leu His Lys Cys Asp Asn Phe Val Gln Ala Val Val Gln Lys Val Lys  
195 200 205  
Glu Asp Ile Thr Pro Ser Ile Glu Val Thr Ile Asp Ser Leu Ile Asp  
210 215 220  
Asp Pro Cys Trp  
225

<210> 20  
<211> 1698  
<212> DNA  
<213> SHRIMP

<400> 20  
atggcaaggt ctgtaggttt attatctgtg acacccgagt atgacacatt taagtatatc 60  
aaaatggaag aattcaagac tttaaaggta aaaaatggat tcactatttc tgggtgaaaac 120  
cctgacaaat acgaacatat tctattatca ttcaaactcg ttgatagggt tacaaaaagt 180  
gaattaaagg acggactata tatagtcgtg ctaaaagata aggaagtact tcacataaaa 240  
aacggaggtgc atagactgag acaattgaca ggcgacaata ctcttcaagt cggattaaag 300  
tatactcaca atctccccag gttagggagt ttattgcaag atgatgggtg tgaggattat 360  
ggggaaaagt ggaacgaatc actccccatt gatatgcaaa atatcaacaa aatcgttaaa 420  
gaaaaggccc ttcttagtga caaaaacttt aaattttctc ctctctacag gcttttacac 480  
gaaagacttt ctaatgcagc tgtgaaaaaa tgtgactata tgataataac cactgacttt 540  
ttagtaggggt gtggctacac gccgagccat tgccctagaa cacttcgtaa catggaacaa 600  
ttattggtgg aacaatgtgg cttttcttcc cgcataatcag tgtatgatat atgtgatagg 660  
ttaacatata aaggggccta cattgcaaac ccaatcacag gcagttactc caacatgtgc 720  
ctaattgttc caatggataa actagggtttg attttctaca acagtacaca cccatcagct 780  
aaaagtattg gaaattatat gtcactctct ttcaatgcaa ccgtcatata cgcaaatgaa 840  
agggataatt tacagatgga taatttcaga agagaaataa agtttgcaga gaatgaagta 900  
aatatgaagg aagaagaact gaaggaattg aggaaacgtt gtgccgtctc tgaagaacaa 960  
aggattttct tgagggatgt gcacaaaaaa tcatcgattg cgacatcccg ctatgatggg 1020

```

ggagcctgct tagtgtttgc tttttctgac cgagatttct ccttgctatg cagaaccaac 1080
ggaaatgggt ccttttattc tgcaacagaa gaaggaatca ggtatgtctc ttcacctgaa 1140
tataaaaaga gagatgtggg agaacgtagg ccccgtttga tcatgtccat aactggatcg 1200
gatgccccta tatgtattag ggacagtgtg cgaaccact tcaagacgag acttttctcc 1260
cgcaccagtg gcaatagtat aaccttttgc gtccctccag gagaaaggga actaatggaa 1320
atgggtcagag aggtaacagg aacagacatc aaaattttca tggataatgg aaaagtatat 1380
caaatgggtg ccgaaatcaa tgtgattgat ccgacatcta aggaatacaa agagttactc 1440
aaaaggggaag aaaacttgcc cgaagacgaa cggaagcggt tgcggcgaga acgacgcatg 1500
attttcaata cgtcaagggc aatttccatg tataatgaag aacgtggaga tggaggaagt 1560
ggaggagaaa cttcagaaga tggagatgga aacggcagca ccagcagtaa aggagagaaa 1620
aggaaaagag aagaaaatga agggaatgaa tatgtagtcc ttttgaacaa ggcttgtaaa 1680
gacataaaag tttgctaa 1698

```

<210> 21  
 <211> 563  
 <212> PRT  
 <213> SHRIMP

<400> 21

Met	Ala	Arg	Ser	Val	Gly	Leu	Leu	Ser	Val	Thr	Pro	Glu	Tyr	Asp	Thr
1				5					10					15	
Phe	Lys	Tyr	Ile	Lys	Met	Glu	Glu	Phe	Lys	Thr	Leu	Lys	Val	Lys	Asn
			20					25					30		
Gly	Phe	Thr	Ile	Ser	Gly	Glu	Asn	Pro	Asp	Lys	Tyr	Glu	His	Ile	Leu
		35					40					45			
Leu	Ser	Phe	Lys	Ser	Val	Asp	Arg	Val	Thr	Lys	Ser	Glu	Leu	Arg	Asp
	50					55					60				
Gly	Ile	Val	Arg	Leu	Lys	Asp	Lys	Glu	Val	Leu	His	Ile	Lys	Asn	Gly
65					70				75					80	
Val	His	Arg	Leu	Arg	Gln	Leu	Thr	Gly	Asp	Asn	Thr	Leu	Gln	Val	Gly
			85						90					95	
Leu	Lys	Tyr	Thr	His	Asn	Leu	Pro	Arg	Leu	Gly	Ser	Leu	Leu	Gln	Asp
			100					105					110		
Asp	Gly	Cys	Glu	Asp	Tyr	Gly	Glu	Lys	Trp	Asn	Glu	Ser	Leu	Pro	Ile
	115						120				125				
Asp	Met	Gln	Asn	Ile	Asn	Lys	Ile	Val	Lys	Glu	Lys	Ala	Leu	Leu	Ser
	130					135					140				
Asp	Lys	Asn	Phe	Lys	Phe	Ser	Pro	Leu	Tyr	Arg	Leu	Leu	His	Glu	Arg
145					150					155					160
Leu	Ser	Asn	Ala	Ala	Val	Lys	Lys	Cys	Asp	Tyr	Met	Ile	Ile	Thr	Thr
			165						170					175	
Asp	Phe	Leu	Val	Gly	Cys	Gly	Tyr	Thr	Pro	Ser	His	Cys	Pro	Arg	Thr
		180						185					190		
Leu	Arg	Asn	Met	Glu	Gln	Leu	Leu	Val	Glu	Gln	Cys	Gly	Phe	Ser	Ser
	195					200					205				
Arg	Ile	Ser	Val	Tyr	Asp	Ile	Cys	Asp	Arg	Leu	Thr	Tyr	Lys	Gly	Ala
	210				215						220				
Tyr	Ile	Ala	Asn	Pro	Ile	Thr	Gly	Ser	Tyr	Ser	Asn	Met	Cys	Leu	Ile
225					230				235					240	
Val	Pro	Met	Asp	Lys	Leu	Gly	Leu	Ile	Phe	Tyr	Asn	Ser	Thr	His	Pro
			245						250					255	
Ser	Ala	Lys	Ser	Ile	Gly	Asn	Tyr	Met	Ser	Ser	Leu	Phe	Asn	Ala	Thr
		260						265					270		
Val	Ile	Tyr	Ala	Asn	Glu	Arg	Asp	Asn	Leu	Gln	Met	Asp	Asn	Phe	Arg
	275						280					285			
Arg	Glu	Ile	Lys	Phe	Ala	Glu	Asn	Glu	Val	Asn	Met	Lys	Glu	Glu	Glu
	290					295					300				
Leu	Lys	Glu	Leu	Arg	Lys	Arg	Cys	Ala	Val	Ser	Glu	Glu	Gln	Arg	Ile
	305				310					315				320	
Ser	Leu	Arg	Asp	Val	His	Lys	Lys	Ser	Ser	Ile	Ala	Thr	Ser	Arg	Tyr
			325						330					335	

Asp Gly Gly Ala Cys Leu Val Phe Ala Phe Ser Asp Arg Asp Phe Ser  
                   340                  345                  350  
 Leu Leu Cys Arg Thr Asn Gly Asn Gly Ser Phe Tyr Ser Ala Thr Glu  
                   355                  360                  365  
 Glu Gly Ile Arg Tyr Val Ser Ser Pro Glu Tyr Lys Lys Arg Asp Val  
                   370                  375                  380  
 Gly Glu Arg Arg Pro Arg Leu Ile Met Ser Ile Thr Gly Ser Asp Ala  
                   385                  390                  395                  400  
 Pro Ile Cys Ile Arg Asp Ser Val Arg Asn His Phe Lys Thr Arg Leu  
                   405                  410                  415  
 Phe Ser Arg Thr Ser Gly Asn Ser Ile Thr Phe Ala Val Pro Pro Gly  
                   420                  425                  430  
 Glu Arg Glu Leu Met Glu Met Val Arg Glu Val Thr Gly Thr Asp Ile  
                   435                  440                  445  
 Lys Ile Phe Met Asp Asn Gly Lys Val Tyr Gln Asn Gly Ala Glu Ile  
                   450                  455                  460  
 Asn Val Ile Asp Pro Thr Ser Lys Glu Tyr Lys Glu Leu Leu Lys Arg  
                   465                  470                  475                  480  
 Glu Glu Asn Leu Pro Glu Asp Glu Arg Lys Arg Leu Arg Arg Glu Arg  
                   485                  490                  495  
 Arg Met Ile Phe Asn Thr Ser Arg Ala Ile Ser Met Tyr Asn Glu Glu  
                   500                  505                  510  
 Arg Gly Asp Gly Gly Ser Gly Gly Glu Thr Ser Glu Asp Gly Asp Gly  
                   515                  520                  525  
 Asn Gly Ser Thr Ser Ser Lys Gly Glu Lys Arg Lys Arg Glu Glu Asn  
                   530                  535                  540  
 Glu Gly Asn Glu Tyr Val Val Leu Leu Asn Lys Ala Cys Lys Asp Ile  
                   545                  550                  555                  560  
 Lys Val Cys

<210> 22  
 <211> 630  
 <212> DNA  
 <213> SHRIMP

<400> 22  
 atgattgtat tcgtcgaagg atctccccta actgggaaaa cttcatgggt ggataacatg 60  
 agaactgccg gaaagggcaa acagtccttc ctcaacttca tgtacaccaa ttatagggac 120  
 tattttacca tcttcccattg gaccattcaa gagcatttgc gtgcatccga ttaccaggaa 180  
 cgtcccaggc tcgtggatgg aatgtttgga tcatcattga actttttcac aggcatgtgg 240  
 aggcacgaca cagaacagtt tccagaatca aaaattggat tgagggagta ttggagatg 300  
 tatggagaag aattcaaagc atgtgtcgcc gagtgggtca aatacaagcc agttttccat 360  
 gtgatggttt acaggggaaga agatgtcaag aaaatggaac caattattca agaattgaat 420  
 gatgcacata actggtttat tgacgttctc aaggaggaga gaggcttgtt tgtaaagatt 480  
 gaagttatcc ccagaaatgt gtacaaaggg aatatttgta gttcatgttt ctcgacgagt 540  
 aagaattacg tgtaccgtgt tggtaaattgc acgaatagta ttgtacactg tgatatgaaa 600  
 tgtaagttaa ttgctgagaa gattatttga 630

<210> 23  
 <211> 209  
 <212> PRT  
 <213> SHRIMP

<400> 23  
 Met Ile Val Phe Val Glu Gly Ser Pro Leu Thr Gly Lys Thr Ser Trp  
   1                  5                  10                  15  
 Val Asp Asn Met Arg Thr Ala Gly Lys Gly Lys Gln Ser Phe Leu Asn  
                   20                  25                  30  
 Phe Met Tyr Thr Asn Tyr Arg Asp Tyr Leu Pro Ile Phe Pro Trp Thr

WO 01/38351

104

PCT/US00/28888

```

      35              40              45
Ile Gln Glu His Leu Arg Ala Ser Asp Tyr Gln Glu Arg Pro Arg Leu
  50              55              60
Val Asp Gly Met Phe Gly Ser Ser Leu Asn Phe Thr Gly Met Trp
  65              70              75              80
Arg His Asp Thr Glu Gln Phe Pro Glu Ser Lys Ile Gly Leu Arg Glu
      85              90              95
Tyr Leu Glu Met Tyr Gly Glu Glu Phe Lys Ala Cys Val Ala Glu Trp
      100              105              110
Val Lys Tyr Lys Pro Val Phe His Val Met Val Tyr Arg Glu Glu Asp
      115              120              125
Val Lys Lys Met Glu Pro Ile Ile Gln Glu Leu Asn Asp Ala His Asn
      130              135              140
Trp Phe Ile Asp Val Leu Lys Glu Glu Arg Ala Leu Phe Val Lys Ile
      145              150              155              160
Glu Val Ile Pro Arg Asn Val Tyr Lys Gly Asn Ile Cys Ser Ser Cys
      165              170              175
Phe Ser Thr Ser Lys Asn Tyr Val Tyr Arg Val Gly Lys Cys Thr Asn
      180              185              190
Ser Ile Val His Cys Asp Met Lys Cys Lys Phe Ile Ala Glu Lys Ile
      195              200              205
Ile

```

<210> 24  
 <211> 1347  
 <212> DNA  
 <213> SHRIMP

```

<400> 24
atgggaggtg aagattcggt cgacgacagg tacgattctg atgcgctctg ggaaaaatgaa 60
ggagcaaaat cgattcaagt aaaggaaacg gatttggagg tgtaccgaat gcaccgcaga 120
gcggtaccca ctctagaaga aaaaaataga acagctttta ggtactattc agatttgtct 180
cctgtttaca gagtgcctct cttctcttta aaggatggaa gtgatcccca cgaaagagat 240
tttagtttca atgtagatcc acgacgattt ggaaaagtac cagtgaaggt gagacgagtt 300
gacgtgagaa atccttcacg aacggctgcc atttctgtgc caactggacc aggacttcat 360
gtatcgtctt acactggaga cgggatgttg gtctgcccta atcacaattt tataggagat 420
ctttgttctg aaatagcatc cgatataaca atttataata caagtagttc agggcggcta 480
tcttatgcaa caaactttta tagtgttgaa gataattccc ctgtcgggat tttattcgag 540
actcttctg acgacaaaaa gttccaacaa gtgtctattt ttcccgctac tgaaccggct 600
tccaatatct ctatcggacc catgagtcac gtgaaaatca agctcgggta ctatgatgaa 660
gagaacgcta ccgcagtcgg tgtcattaga tatggaggat tattctacac ctctgtcggg 720
gcatgcatta tcccagaagg ggttttcttt gatgatgttg tgggaaatca tagctcaatg 780
aacatataca acatgactaa ccaacaaaaa gaaattgtac taaaagagcc aagaggagag 840
gatgccatgg aagaagatga tgggtgaagaa gcagattata atttcttggg ttatgtagta 900
agatttgaac atgatttgaa aatgcaagcc atgtcgtcag catatagctc agtttctatt 960
gatattaaca gttcaagttt ccacaaatgt ttctaatta aacctaaata taactctatt 1020
ttacaacctc tcgtttcttc agaagttgtg cttaatgata tctcgtctca tacacgtggg 1080
agagaagtag agtttcatga caggctcccc tctgggtcac aagacaactc atattcgatt 1140
gtaaaatata tgaaaatagt atctctaaag gaaggcctaa aggttgtaaa cctataatc 1200
aatacagaat tgtataagaa aaaacaggct ttaaagggtg atgtttttaa tatgacgcgt 1260
gatgttggtg gactagatac gagtgcagat agttttggag taatagtgtg tcatgctgca 1320
aaacttcccc aagtgattgg tcaataa 1347

```

<210> 25  
 <211> 444  
 <212> PRT  
 <213> SHRIMP

<400> 25

```

Met Gly Gly Glu Asp Ser Phe Asp Asp Arg Tyr Asp Ser Asp Ala Leu
 1      5      10
Trp Glu Asn Glu Gly Ala Lys Ser Ile Gln Val Lys Glu Thr Asp Leu
      20      25      30
Glu Val Tyr Arg Met His Arg Arg Ala Val Pro Thr Leu Glu Glu Lys
      35      40      45
Asn Arg Thr Ala Leu Arg Tyr Tyr Ser Asp Trp Ser Pro Val Tyr Arg
      50      55      60
Val Pro Leu Phe Ser Leu Lys Asp Gly Ser Asp Phe Arg Asp Phe Ser
65      70      75      80
Phe Asn Val Asp Pro Arg Arg Phe Gly Lys Val Pro Val Lys Val Arg
      85      90      95
Arg Val Asp Val Arg Asn Pro Ser Arg Thr Ala Ala Ile Phe Val Pro
      100      105      110
Thr Gly Pro Gly Leu His Val Ser Ser Tyr Thr Gly Asp Gly Met Leu
      115      120      125
Val Cys Pro Asn His Asn Phe Ile Gly Asp Leu Cys Ser Glu Ile Ala
      130      135      140
Ser Asp Ile Thr Ile Tyr Asn Thr Ser Ser Ser Gly Arg Leu Ser Tyr
145      150      155      160
Ala Thr Asn Phe Asn Ser Val Glu Asp Asn Ser Pro Val Gly Ile Leu
      165      170      175
Phe Glu Thr Leu Pro Asp Asp Lys Met Phe Gln Gln Val Ser Ile Phe
      180      185      190
Ser Ala Thr Glu Pro Asn Ile Ser Ile Gly Pro Met Ser His Val Lys
      195      200      205
Ile Lys Leu Gly Tyr Tyr Asp Glu Glu Asn Ala Thr Ala Val Gly Val
      210      215      220
Ile Arg Tyr Gly Gly Leu Phe Tyr Thr Ser Val Gly Ala Cys Ile Ile
225      230      235      240
Pro Glu Gly Val Phe Phe Asp Asp Val Val Gly Asn His Ser Ser Met
      245      250      255
Asn Ile Tyr Asn Met Thr Asn Gln Pro Lys Glu Ile Val Leu Lys Glu
      260      265      270
Pro Arg Gly Glu Asp Ala Met Glu Glu Asp Asp Gly Glu Glu Ala Asp
      275      280      285
Tyr Asn Phe Leu Gly Tyr Val Arg Phe Glu His Asp Leu Lys Met
      290      295      300
Gln Ala Met Ser Ser Ala Tyr Ser Ser Val Ser Ile Asp Ile Asn Ser
305      310      315      320
Ser Ser Phe His Lys Cys Phe Leu Ile Lys Pro Lys Tyr Asn Ser Ile
      325      330      335
Leu Gln Pro Leu Val Ser Ser Glu Val Val Leu Asn Asp Leu Ser Leu
      340      345      350
Asn Thr Arg Gly Arg Glu Val Glu Phe His Asp Arg Leu Pro Ser Gly
      355      360      365
Ala Gln Asp Asn Ser Tyr Ser Ile Val Lys Tyr Met Lys Ile Val Ser
      370      375      380
Leu Lys Glu Gly Leu Lys Val Val Asn Pro Ile Ile Asn Thr Glu Leu
385      390      395      400
Tyr Lys Lys Lys Gln Ala Leu Lys Val His Val Leu Asn Met Thr Arg
      405      410      415
Asp Val Val Gly Leu Asp Thr Ser Glu His Ser Phe Gly Val Ile Val
      420      425      430
Cys His Ala Ala Lys Leu Pro Glu Val Ile Gly Gln
      435      440

```

&lt;210&gt; 26

&lt;211&gt; 2352

&lt;212&gt; DNA

<213> SHRIMP

<400> 26

```

atggcagcag cagcagtctc aggagagggg agaatctctg cagatctact cctgttgga 60
caactcaccc cggacggaga cgtgataaga tacgactctg agcagtacac caaacctagg 120
aagatctttg gtgacaaaag tgtgatagag actattggac attttctcat ccacaaccac 180
aaccaagggt agagttacca aatcgcatct tctgtgttg aaaaattccc cgctctactc 240
aattgcatat ggaatggaga gtcgggagga atggctctat ggaaggcatt gtacagggct 300
aaaaagtata gacttctcaa ctcacttttg gtccacaaga taaaaaattg gccttcgggt 360
gccgtgatcc ctatctacgg gtcagtgtgt gacagagaag aaaggcccat catcatgagt 420
gagattattg acaaggaac tcttcagacc atatgtaaga gtgatatacg ttctcttctt 480
ggaatgatga acgccaagca tggcacattg ggaggttaatt ttttacactt ctatgcccg 540
tcaactaaac cgttgaaaaa tttccaatat gaagcaatgg gagctaattg agtgctaattg 600
gcagctgaag ctatttatga tggattcaga gaccatggct taaaccatc agaataact 660
tttcttggtt tggaatctgc tgatgtgtac ggaacaatc cagtggaaat tgcaatatca 720
ggagatgatg acaatatggt gttgaacctc atctgcaact atgggtgtatc ttatgaaaaa 780
actcgcggtc gggtaaatag atccttggtt gattttttta aaatgaacac agcttcaaa 840
tgtctcagtg ttttaaagtt tgttgaaaaa cactttaaaa ttgagtcaaa tacacaaaa 900
ggagaatttg aagaaaaggc cgaaacatgt gttaattgtc ttgatagaaa taatgtgttg 960
actaaagggt ctgaacaaga atcttacaag ctttcttggt gacactttct tcatgtaaaa 1020
tgtttgagga atatttgat agtatcacia cacctaagat gtgaaaaatg tctaaaaaga 1080
tttgatgaga gtattttgag aaagtgtaca cctaacttga attgggtggt gactatgccg 1140
gcaggtgctg gaaatgaaga agaaatatgt ttcatgagaa ataagaaact ggttgatgat 1200
ttcagaaaat tgttgtcccc tgtctcaatt cctcatttct tcaaaaatag tagacagcga 1260
aatcttgata tgttgtgccc ctacagtgtc cacacgataa taccaaataa agaagatcca 1320
aagaaaaacg aagatggaac cagagtgtgg gtcaaccaca cagccatcag tgaaaagcag 1380
aacaaggagg aagaagacgc gaggataaag cgtgtagccg tcaggacatt tacagccatc 1440
agagaaaagc agaacaagga ggaagaagac gcgaggatca agcgtgcagt cgacatggct 1500
gtcgagccca tcaacgaaaa gaacaaggag gaagaagacg cgaggatcaa gcgtgcagtc 1560
gacatggctg tcgcagccat caacgaaaat aacaaggagg aagaagacgc gaggatcaag 1620
cgtgcagctg acatggcggt tgcagccatc aacgaaaaga acaaggagga agaagacgcg 1680
aggatcaagc gtgcagtcga catggctgtc gcagccatca acgaaaataa caaggaggaa 1740
gaagacgcga ggatcaagcg tgcagtcgac atggcggttg cagccatcaa cgaaaataac 1800
aaggaggaag aagacgcgag gatcaagcgt gcagtcgaca tggcggttgc agccaccaac 1860
gaaaagaaca agaaggaaga agacgcgagg atcaagcgta taattgactt gactgttgat 1920
atgaggattc aacgtatagt cgacatggca attgcagctg ccactaaaaa ggacaagaaa 1980
gaagaagaga aaaggacaaa aagggaacaa gagttaaggg ctgatctgag aagggcaatg 2040
gatatggtga acgaagtaca gaagaaactt gaagacatgg aactagaaaa ggggtgtaat 2100
aaggatgaag ccaagaatac tagtaattgt gttagtagca gtagtgttgt tgcctattct 2160
aaagaaattg taccttggtt aggaataat aataatgctg tcattggtat gactagcacc 2220
aactattctg ccaacaatac taagaataat gtatttggtt cacctcataa attttcttc 2280
aacgatgcat ctagattctc caatattgta gaaactccca aaatgtcttt caatttctcg 2340
ttcaagacat aa 2352

```

<210> 27

<211> 781

<212> PRT

<213> SHRIMP

<400> 27

```

Met Ala Ala Ala Val Ser Gly Glu Gly Arg Ile Ser Ala Asp Leu
 1          5          10          15
Leu Leu Leu Glu Gln Leu Thr Pro Asp Gly Asp Val Ile Arg Tyr Asp
          20          25          30
Ser Glu Gln Tyr Thr Lys Pro Arg Lys Ile Phe Gly Asp Lys Ser Val
          35          40          45
Ile Glu Thr Ile Gly His Phe Leu Ile His Asn His Asn Gln Gly Glu
          50          55          60
Ser Tyr Gln Ile Ala Ser Ser Val Leu Glu Lys Phe Pro Ala Leu Leu
          65          70          75          80
Asn Cys Ile Trp Asn Gly Glu Ser Gly Gly Met Ala Leu Trp Lys Ala

```

Leu	Tyr	Arg	Ala	85				Leu	90				Leu	Leu	Val	His
				Lys	Lys	Tyr	Arg		Leu	Leu	Asn	Ser				
Lys	Ile	Lys	100	Trp	Pro	Ser	Val	105	Ala	Val	Ile	Pro	Ile	Tyr	Gly	Ser
			Asn					120								
Val	Cys	Asp	Arg	Glu	Glu	Arg	Pro	Ile	Ile	Met	Ser	Glu	Ile	Ile	Asp	
																130
Lys	Glu	Thr	Leu	Gln	Thr	Ile	Cys	Lys	Ser	Asp	Ile	Arg	Ser	Leu	Leu	
																145
Gly	Met	Met	Asn	Ala	Lys	His	Gly	Thr	Leu	Gly	Gly	Asn	Phe	Leu	His	
																165
Phe	Tyr	Ala	Arg	Ser	Thr	Lys	Pro	Phe	Glu	Asn	Phe	Gln	Tyr	Glu	Ala	
																180
Met	Gly	Ala	Asn	Ala	Val	Leu	Met	Ala	Ala	Glu	Ala	Ile	Tyr	Asp	Gly	
																195
Phe	Arg	Asp	His	Gln	Pro	Ser	Glu	Tyr	Thr	Phe	Pro	Gly	Leu	Glu	Ser	
																210
Ala	Asp	Val	Tyr	Gly	Asn	Asn	Pro	Val	Glu	Ile	Ala	Ile	Ser	Gly	Asp	
																225
Asp	Asp	Asn	Met	Leu	Asn	Leu	Ile	Cys	Asn	Tyr	Gly	Val	Ser	Tyr	His	
																245
Glu	Lys	Thr	Arg	Gly	Arg	Val	Asn	Arg	Ser	Leu	Leu	Asp	Phe	Leu	Lys	
																260
Met	Asn	Thr	Ala	Ser	Lys	Cys	Leu	Ser	Val	Leu	Lys	Phe	Val	Glu	Lys	
																275
His	Phe	Lys	Ile	Glu	Ser	Asn	Thr	Pro	Lys	Gly	Glu	Phe	Glu	Glu	Lys	
																290
Ala	Glu	Thr	Cys	Val	Asn	Cys	Leu	Asp	Arg	Asn	Asn	Val	Leu	Thr	Lys	
																305
Gly	Ser	Glu	Gln	Glu	Ser	Tyr	Lys	Leu	Ser	Cys	Gly	His	Phe	Leu	His	
																325
Val	Lys	Cys	Leu	Arg	Asn	Ile	Cys	Ile	Val	Ser	Gln	His	Leu	Arg	Cys	
																340
Glu	Lys	Cys	Leu	Lys	Arg	Phe	Asp	Glu	Ser	Ile	Leu	Arg	Lys	Cys	Thr	
																355
Pro	Asn	Leu	Asn	Trp	Trp	Leu	Thr	Met	Pro	Ala	Gly	Ala	Gly	Asn	Glu	
																370
Glu	Glu	Ile	Cys	Phe	Met	Arg	Asn	Lys	Lys	Leu	Val	Asp	Asp	Phe	Arg	
																385
Lys	Leu	Leu	Ser	Pro	Val	Ser	Ile	Pro	His	Phe	Phe	Lys	Asn	Ser	Arg	
																405
Gln	Arg	Asn	Leu	Asp	Met	Leu	Cys	Pro	Tyr	Ser	Asp	His	Thr	Ile	Ile	
																420
Pro	Asn	Lys	Glu	Asp	Pro	Lys	Lys	Asn	Glu	Asp	Gly	Asn	Arg	Val	Arg	
																435
Val	Asn	His	Thr	Ala	Ile	Ser	Glu	Lys	Gln	Asn	Lys	Glu	Glu	Asp	His	
																450
Ala	Arg	Ile	Lys	Arg	Val	Ala	Val	Arg	Thr	Phe	Thr	Ala	Ile	Arg	Glu	
																465
Lys	Gln	Asn	Lys	Glu	Glu	Glu	Asp	Ala	Arg	Ile	Lys	Arg	Ala	Val	Asp	
																485
Met	Ala	Val	Ala	Ala	Ile	Asn	Glu	Lys	Asn	Lys	Glu	Glu	Glu	Asp	Ala	
																500
Arg	Ile	Lys	Arg	Ala	Val	Asp	Met	Ala	Val	Ala	Ala	Ile	Asn	Glu	Asn	
																515
Asn	Lys	Glu	Glu	Glu	Asp	Ala	Arg	Ile	Lys	Arg	Ala	Val	Asp	Met	Ala	
																530
Val	Ala	Ala	Ile	Asn	Glu	Lys	Asn	Lys	Glu	Glu	Glu	Asp	Ala	Arg	Ile	
																545
Lys	Arg	Ala	Val	Asp	Met	Ala	Val	Ala	Ala	Ile	Asn	Glu	Asn	Asn	Lys	
																565

Glu Glu Glu Asp Ala Arg Ile Lys Arg Ala Val Asp Met Ala Val Ala  
 580 585 590  
 Ala Ile Asn Glu Asn Asn Lys Glu Glu Glu Asp Ala Arg Ile Lys Arg  
 595 600 605  
 Ala Val Asp Met Ala Val Ala Thr Asn Glu Lys Asn Lys Lys Glu  
 610 615 620  
 Glu Asp Ala Arg Ile Lys Arg Ile Ile Asp Leu Thr Val Asp Met Arg  
 625 630 635 640  
 Ile Gln Arg Ile Val Asp Met Ala Ile Ala Ala Thr Lys Lys Asp  
 645 650 655  
 Lys Lys Glu Glu Glu Lys Arg Thr Lys Arg Glu Gln Glu Leu Arg Ala  
 660 665 670  
 Asp Leu Arg Arg Ala Met Asp Met Val Asn Glu Val Gln Lys Lys Leu  
 675 680 685  
 Glu Asp Met Glu Leu Glu Lys Gly Cys Asn Lys Asp Glu Ala Lys Asn  
 690 695 700  
 Thr Ser Asn Val Val Ser Ser Ser Ser Val Val Ala Tyr Ser Lys Glu  
 705 710 715 720  
 Ile Val Pro Cys Leu Gly Asn Asn Asn Asn Ala Val Ile Gly Met Thr  
 725 730 735  
 Ser Thr Asn Tyr Ser Ala Asn Asn Thr Lys Asn Asn Val Phe Gly Ser  
 740 745 750  
 Pro His Lys Phe Ser Phe Asn Asp Ala Ser Arg Phe Ser Asn Ile Val  
 755 760 765  
 Glu Thr Pro Lys Met Ser Phe Asn Phe Ser Phe Lys Thr  
 770 775 780

&lt;210&gt; 28

&lt;211&gt; 1506

&lt;212&gt; DNA

&lt;213&gt; SHRIMP

&lt;400&gt; 28

atggtatata aaggttttgt atctccatca tttatcacta ttcgtacgat gacgtccaac 60  
 agaccaacaa catccctctt ttcgttctct gaaggttctt cgctttcagg ggataagtat 120  
 gatacgtatg aggatattctt acttgaacaa ttcaactgtt tcaagacatc ttctccttct 180  
 tctgtctgta aaagtgaat agaggataaa actttaatct ttcaacttaa agaaggagaa 240  
 aaattccatc ttgcaaagg catagaagag ctccgtgaga ttctagacga taattctgca 300  
 acaattgaac ctattatttc tccaacgaca ttcaatgaca gaaacgaatt actaaaccac 360  
 gagggagata tatcctcaag tcccctatat actcagataa tgaagcatat ttcaccagag 420  
 catgatattt atgaattgga cttatttgtt ggcactgatt tgctttttgg tctagggtgtg 480  
 aatctacgca acgtttctaa actgatgaag aaaatatcgt atggtacttt aaatgtagtt 540  
 gatgtgtgcc acagaaaatt tttcaacaat aggattatag ttaatcccat ttcttcgtca 600  
 ttctcaaaaa atgtgtgtat tattcctcta ttttctgcag ctgaagaatt ttcgtctctg 660  
 ggggaatgca gggatttatt caacggtatt tgtgatgacg tagagagata tatcaactct 720  
 tattttttct accctgaaaa tactactact actactacta ctgctccttc gtcgcccga 780  
 atggaaattg cagatgagga agaacaatcc ccaaaaacta taaagagaaa tgacaacgca 840  
 agtagaaact ggtctggtgt ctgtttgatt tttgaagtat ttaaaaacac gtactacatt 900  
 attaatagag gagatagagg aggttctttt gaaaaggctg tgaagagtgc aatttcttct 960  
 atcaaggaaa agagatgtaa aatcacagat attaattggt ataaacctcg attggttatg 1020  
 gtgataactg ggtgtttatc agaattgtac ttcaaagatg cacttaaaca gattggagaa 1080  
 aacaggcgca aatttttgaa aatgaatggg aattactttt ctctgattga tgaacaagca 1140  
 gatctaactc aattcgcgat gagtgtttct ggtgccgggg agaggatttt tggttaacggt 1200  
 ttggggatgt tccagaaccg taaaatgata cctgtaattg atcctctcac atatgaaaat 1260  
 gttgtatgtg gtgagcatga tatacaaaaa gaagatgcta ttctttctgt aaggagagct 1320  
 attgcagact ataatgactt tgtaagtaag aacaagagag ggaagaaaac cagcgagaa 1380  
 gaagaaaatg aagatgaaga tgcagacgct agcagcagca gcagcagcag tctcctcct 1440  
 tcttctcctc ctgcacataa aaaatcacgt cttccggatg aaggcgaaaa atgtacactc 1500  
 tggttaa 1506



<210> 29  
 <211> 499  
 <212> PRT  
 <213> SHRIMP

<400> 29

Met	Val	Tyr	Lys	Gly	Phe	Val	Ser	Pro	Ser	Phe	Ile	Thr	Ile	Arg	Thr
1				5					10					15	
Met	Thr	Ser	Asn	Arg	Pro	Thr	Thr	Ser	Pro	Leu	Ser	Phe	Ser	Glu	Gly
			20					25					30		
Phe	Ser	Leu	Ser	Gly	Asp	Lys	Tyr	Asp	Thr	Tyr	Glu	Asp	Ile	Leu	Leu
		35					40					45			
Glu	Gln	Phe	Asn	Cys	Phe	Lys	Thr	Ser	Ser	Pro	Ser	Ser	Ala	Arg	Lys
	50					55					60				
Ser	Glu	Ile	Glu	Asp	Lys	Thr	Leu	Ile	Phe	Gln	Leu	Lys	Glu	Gly	Glu
65					70					75					80
Lys	Phe	His	Leu	Ala	Lys	Gly	Ile	Glu	Glu	Leu	Arg	Glu	Ile	Leu	Asp
			85						90					95	
Asp	Asn	Ser	Ala	Thr	Ile	Glu	Pro	Ile	Ile	Ser	Pro	Thr	Thr	Phe	Asn
			100					105					110		
Asp	Arg	Asn	Glu	Leu	Leu	Asn	His	Glu	Gly	Asp	Ile	Ser	Ser	Ser	Pro
		115					120					125			
Leu	Tyr	Thr	Gln	Ile	Met	Lys	His	Pro	Glu	His	Asp	Ile	Tyr	Glu	Leu
	130					135					140				
Asp	Leu	Ile	Val	Gly	Thr	Asp	Leu	Leu	Phe	Gly	Leu	Gly	Val	Asn	Leu
145					150					155					160
Arg	Asn	Val	Ser	Lys	Leu	Met	Lys	Lys	Ile	Ser	Tyr	Gly	Thr	Leu	Asn
				165					170					175	
Val	Val	Asp	Val	Cys	His	Arg	Lys	Phe	Phe	Asn	Asn	Arg	Ile	Ile	Val
			180					185					190		
Asn	Pro	Ile	Ser	Ser	Ser	Phe	Ser	Lys	Asn	Val	Cys	Ile	Ile	Pro	Leu
		195					200					205			
Phe	Ser	Ala	Ala	Glu	Glu	Phe	Ser	Ser	Leu	Gly	Glu	Cys	Arg	Asp	Leu
		210				215					220				
Phe	Asn	Gly	Ile	Cys	Asp	Asp	Val	Glu	Arg	Tyr	Ile	Asn	Ser	Tyr	Phe
225					230					235					240
Phe	Tyr	Pro	Glu	Asn	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Ala	Pro	Ser	Ser
				245					250					255	
Pro	Glu	Met	Glu	Ile	Ala	Asp	Glu	Glu	Glu	Gln	Ser	Pro	Lys	Thr	Ile
			260					265					270		
Lys	Arg	Asn	Asp	Asn	Ala	Ser	Arg	Asn	Trp	Ser	Gly	Val	Cys	Leu	Ile
		275					280					285			
Phe	Glu	Val	Phe	Lys	Asn	Thr	Tyr	Tyr	Ile	Ile	Asn	Arg	Gly	Asp	Arg
	290					295					300				
Gly	Gly	Ser	Phe	Glu	Lys	Ala	Val	Lys	Ser	Ala	Ile	Ser	Ser	Ile	Lys
305					310					315					320
Glu	Lys	Arg	Cys	Lys	Ile	Thr	Asp	Ile	Asn	Gly	Asn	Lys	Pro	Arg	Leu
				325					330					335	
Val	Met	Val	Ile	Thr	Gly	Cys	Tyr	Thr	Glu	Leu	Tyr	Phe	Lys	Asp	Ala
			340					345					350		
Leu	Lys	Gln	Ile	Gly	Glu	Asn	Arg	Arg	Lys	Phe	Leu	Lys	Met	Asn	Gly
		355					360					365			
Asn	Tyr	Phe	Ser	Leu	Ile	Asp	Glu	Gln	Ala	Asp	Leu	Ile	Glu	Phe	Ala
	370					375					380				
Met	Ser	Val	Ser	Gly	Ala	Gly	Glu	Arg	Ile	Phe	Val	Asn	Gly	Leu	Gly
385					390					395					400
Met	Phe	Gln	Asn	Arg	Lys	Met	Ile	Pro	Val	Ile	Asp	Pro	Leu	Thr	Tyr
				405					410					415	
Glu	Asn	Val	Val	Cys	Gly	Glu	His	Asp	Ile	Gln	Lys	Glu	Asp	Ala	Ile
			420					425					430		
Leu	Ser	Val	Arg	Arg	Ala	Ile	Ala	Asp	Tyr	Asn	Asp	Phe	Val	Ser	Lys

435                      440                      445  
 Asn Lys Arg Gly Lys Lys Arg Ser Ala Glu Glu Glu Asn Glu Asp Glu  
 450                      455                      460  
 Asp Ala Asp Ala Ser Ser Ser Ser Ser Ser Pro Pro Pro Ser Ser  
 465                      470                      475                      480  
 Pro Pro Ala His Lys Lys Ser Arg Leu Pro Asp Glu Gly Glu Lys Cys  
 485                      490                      495  
 Thr Leu Cys

<210> 30  
 <211> 846  
 <212> DNA  
 <213> SHRIMP

<400> 30  
 atggcggtaa acttggataa tgttcttgtg aatatcaaca acaaggatga agatcttaca 60  
 aaactcgtat ccgaggcaat aaagcggcga gctaaaactg tatttgacac taaaaatcaa 120  
 gcagggtttg acatgagacg tcaagttgaa gctgcattat atgaagcaat atccaaaaag 180  
 aaagaaaagg ccataaaggc attcgatgag ctcatacaag aaagaggtga tgaaattaca 240  
 cctttgacta caatgcagta tgaagagtgg gtaaaccgta caataactcc ctcatgacg 300  
 actgaaaatt tattaggtga tgttgagcac gccgattttt tactggaccg aatgacaccc 360  
 gtaagcgaag aagatattga aggtttcgct gcttctactt ttaaggaggt atcagattca 420  
 aaaactgcaa cagtcatagt taaggcagat tgtgaaacgg gggatatcga tgaagtgtat 480  
 aatcttgac catcattcgg cgtcactcaa gaaattaaaa tatataggtc aaacaattcc 540  
 tcggaattgg ataattgtcg agattctttc catatttata aaatttctgc aacagatagc 600  
 gacagtggca atactaaaaa attgttgtat ggggttaagga ataaaaaagc aggtttatagc 660  
 tgtttgtgta gaatttttgc agaaattgaa tcagatggga ttatggccaa tacaaatatc 720  
 ggtgtcgtcg aaaacaacag agatgaaatt gatgaaaacg aagaaggtaa atatggtttt 780  
 ttaataccca aacaaccagc tgggtcaaaa ttgatcatct acttcttttt aaattgtttg 840  
 acataa 846

<210> 31  
 <211> 282  
 <212> PRT  
 <213> SHRIMP

<220>  
 <221> VARIANT  
 <222> (1)...(282)  
 <223> Xaa = Any Amino Acid

<400> 31  
 Met Ala Val Asn Leu Asp Asn Val Leu Val Asn Ile Asn Asn Lys Asp  
 1                      5                      10                      15  
 Glu Asp Leu Thr Lys Leu Val Ser Glu Ala Ile Lys Arg Arg Ala Lys  
 20                      25                      30  
 Thr Val Phe Asp Thr Lys Asn Gln Ala Gly Phe Asp Met Arg Arg Gln  
 35                      40                      45  
 Val Glu Ala Ala Leu Tyr Glu Ala Ile Ser Lys Lys Lys Glu Lys Ala  
 50                      55                      60  
 Ile Lys Ala Phe Asp Glu Leu Ile Gln Glu Arg Gly Asp Glu Ile Thr  
 65                      70                      75                      80  
 Pro Leu Thr Thr Met Gln Tyr Glu Glu Trp Val Asn Arg Thr Ile Thr  
 85                      90                      95  
 Pro Ser Leu Thr Thr Glu Asn Leu Leu Gly Asp Val Glu His Ala Asp  
 100                      105                      110  
 Phe Leu Leu Asp Arg Met Thr Pro Val Ser Glu Glu Asp Ile Glu Gly  
 115                      120                      125  
 Phe Ala Ala Ser Thr Phe Lys Glu Val Ser Asp Ser Lys Thr Ala Thr

<400> 33															
Met	Thr	Val	Leu	Ala	Val	Tyr	Thr	Ala	Pro	Gln	Ile	Lys	Lys	Ser	Lys
1				5					10					15	
Lys	Arg	Lys	Ile	Glu	Asp	Glu	Asn	Glu	Glu	Pro	Val	Lys	Thr	Leu	
			20					25					30		
Glu	Asp	Phe	Val	Lys	Gly	Arg	Leu	Leu	Asn	Ala	Val	Lys	Glu	Lys	Pro
		35					40					45			
Ala	Glu	Tyr	Phe	Glu	Leu	Leu	Ile	Ser	Ala	Asp	Thr	Glu	Ala	Ala	Leu
	50					55					60				

Lys Thr Ala Glu Glu Thr Ala Leu Arg Asp Phe Val Ile Glu Asn Asp  
 65 70 75 80  
 Ser Val Glu Ile Asp Val Glu Glu Val Leu Glu Glu Lys Pro Arg Glu  
 85 90 95  
 Tyr Val Phe Lys Leu Ala Gly Ala Thr Ser Glu Thr Leu Thr Asn Thr  
 100 105 110  
 Ile Ile Ala Glu Val Gln Lys Lys Ala Ala Leu Ile Thr Glu Glu Asp  
 115 120 125  
 Ile Thr Ile Lys Met Leu Lys Gln Phe Arg Ala Ala Asn Lys Asp Asn  
 130 135 140  
 Lys Asp Gly Glu Ala Thr Pro Glu Glu Lys Glu Asp Phe Thr Asn Asn  
 145 150 155 160  
 Ser Asp Leu Val Gly Leu Asn Glu Val Val Glu Lys Thr Thr Asn Ile  
 165 170 175  
 Val Ile Asn Lys Ile Phe Phe Met Val Phe Glu Arg Cys Ala Ile Leu  
 180 185 190  
 Ile Glu Asp Phe Asp Thr Gly Val Val Thr Asp Gln Ala Ile Gln Ile  
 195 200 205  
 Pro Ser Asn Lys Tyr Lys Ile Arg Leu Val Glu Gly Asp Glu Pro Glu  
 210 215 220  
 Val Phe Pro Gly Asp Cys Leu Asp Leu Ala Val Ser Val Asp Lys Ile  
 225 230 235 240  
 Asn His Val Leu Lys Ile Ser Ala Lys Asn Gly Cys Glu Asn Asn Cys  
 245 250 255  
 Phe Val Ile Ile Pro Arg Phe Ser Pro Val Gly Ser Val Ser Ser Met  
 260 265 270  
 Ile Leu Gly Ser Thr Asp Gln Val Lys Pro Lys Thr Phe Leu Phe Leu  
 275 280 285  
 Ala Asn Lys Asn Asp Ser Thr His Phe Gln Phe Thr Met Asp Lys Gln  
 290 295 300  
 His Ser Val Gly Cys Glu Leu Asp Met Leu Ile Phe Ser Asn Leu Arg  
 305 310 315 320  
 Asn Leu Pro Asp Ser Lys Pro Arg Pro Leu Ser Asp Ala Asp Ile Leu  
 325 330 335  
 Ala Ser Tyr Gly Lys Arg Leu Gly Thr Gly Val Phe Thr Thr Glu Asn  
 340 345 350  
 Leu Val Asp Asp  
 355

&lt;210&gt; 34

&lt;211&gt; 930

&lt;212&gt; DNA

&lt;213&gt; SHRIMP

&lt;400&gt; 34

atgtcttctt cgtcttctga aactcctaag acttcacacg atactgggga agaaaggatt 60  
 aaagacattg taaatgctct agataataat ggcgagtggg tgtcttccta tattgatccg 120  
 attatcaata atcacatctc acgaaaaacg gcagaaactg tccaaaaaat caaccaagaa 180  
 gttgatgaac ggtacgatag aaaaatagcc gacaaaatca acgaaataaa atcatccatc 240  
 ttacaaagtg ctcagactat gtatgaccaa tatgcaatag acacatttca agaaggaaaa 300  
 ggagccaacg ggactggacc agtcatgggg ccagtgaaca cggttatcga tacaacttta 360  
 aataaaatga ggggaaatat gctcgaatac gctgaagata tgtgggacgg agatgactgg 420  
 aaacgatttt ccagttctat gacaacgctt gaatttgatc taagttactc tgatttaact 480  
 atgatgcgtg gttctgacgg gtattttgca ttccttttcc gtggaacaaa aaagataaag 540  
 atggacggtt caagaaagaa agaagaccca attaattgta tcatttcagt aacatatcca 600  
 aacaaagtag gggatgagtg ggaagagggt aaagaacgtg aagtgaattt taacctagaa 660  
 agagtagacg actatgaaag agatatccat gtttcaattt tgtgcatgtt acatgcacaa 720  
 cttgataatt tcgaacaagc attaggagaa aatgcaaaact ctttttattt taaaaagggg 780  
 caaagagtca tgttcttacc caagaaatct aaactgttca atagacctac tgtagaagat 840  
 tctgatatgt tttctataat attcccacct gcacttgacc aagattttgc agatgatatt 900

tattatcgaa taattgtaac atgttcataa

930

<210> 35  
 <211> 307  
 <212> PRT  
 <213> SHRIMP

<400> 35  
 Met Ser Ser Ser Ser Ser Glu Thr Pro Lys Thr Ser Thr Asp Thr Gly  
 1 5 10 15  
 Glu Glu Arg Ile Lys Asp Ile Val Asn Ala Leu Asp Asn Asn Gly Glu  
 20 25 30  
 Trp Leu Ser Ser Tyr Ile Asp Pro Ile Ile Asn Asn His Arg Lys Thr  
 35 40 45  
 Ala Glu Thr Val Gln Lys Ile Asn Gln Glu Val Asp Glu Arg Tyr Asp  
 50 55 60  
 Arg Lys Ile Ala Asp Lys Ile Asn Glu Ile Lys Ser Ser Ile Phe Thr  
 65 70 75 80  
 Ser Ala Gln Thr Met Tyr Asp Gln Tyr Ala Ile Asp Thr Phe Gln Glu  
 85 90 95  
 Gly Lys Gly Ala Asn Gly Thr Gly Pro Val Met Gly Pro Val Asn Thr  
 100 105 110  
 Val Ile Asp Thr Thr Leu Asn Lys Met Arg Gly Asn Met Leu Glu Tyr  
 115 120 125  
 Ala Glu Asp Met Trp Asp Gly Asp Asp Trp Lys Arg Phe Ser Ser Ser  
 130 135 140  
 Met Thr Thr Leu Glu Phe Asp Leu Ser Tyr Ser Asp Leu Thr Met Met  
 145 150 155 160  
 Arg Gly Ser Asp Gly Tyr Phe Ala Phe Pro Phe Arg Gly Thr Lys Lys  
 165 170 175  
 Ile Lys Met Asp Gly Ser Arg Lys Lys Glu Asp Pro Ile Asn Cys Ile  
 180 185 190  
 Ile Ser Val Thr Tyr Pro Asn Lys Val Gly Asp Glu Trp Glu Glu Gly  
 195 200 205  
 Lys Glu Arg Glu Val Asn Phe Asn Leu Glu Arg Val Asp Asp Tyr Glu  
 210 215 220  
 Arg Asp Ile His Val Ser Ile Leu Cys Met Leu His Ala Gln Leu Asp  
 225 230 235 240  
 Asn Phe Glu Gln Ala Leu Gly Glu Asn Ala Asn Ser Phe Tyr Phe Lys  
 245 250 255  
 Lys Gly Gln Arg Val Met Phe Leu Pro Lys Lys Ser Lys Leu Phe Asn  
 260 265 270  
 Arg Pro Thr Val Glu Asp Ser Asp Met Phe Ser Ile Ile Phe Pro Pro  
 275 280 285  
 Ala Ser Asp Gln Asp Phe Ala Asp Asp Ile Tyr Tyr Arg Ile Ile Val  
 290 295 300  
 Thr Cys Ser  
 305

<210> 36  
 <211> 3657  
 <212> DNA  
 <213> SHRIMP

<400> 36  
 atggagacaa ctatggataa tgtcgttcag aataacgacg taacaaaacc aacaccagat 60  
 gttgctactg ttacaacagc aactgagaaa cgtcagtcac gcaaagagaa aaaggatcaa 120  
 cttaaggccg aatgtcctca agtactgaga gcactaaaat tgtccaatac tttaaaggca 180  
 aattttgtaa aatccatgtc ggctatTTTT gctcaacatt tagtggacat gacaaacgct 240  
 aaacacttta aggacccaaa gacaaagaag attttagaac tggatggaag tagtagcagt 300

```

gacagtgaag aagaggaaga aactagttct tcatccaaac ggaaaagagg tagtggtgct 360
agaagtgcct cttcaaagaa agaaaaatgc cccaatacta tcaaaaattg gctcaatgat 420
gctcaagggtg tattccgcga gtttgagat atcatcatta atcttccctc ttttgatgat 480
cttagagacg aagtaaagga tgaacaaact gagctaaaga ccatatatga cttgtataga 540
caggacatgg aaaaggtggt ggaagaagtt ttagggcgcc aagacctgtt tgatcacaag 600
tcagaaatag ccaaaggttt ggcccgtttc gataccacg tctcgttgct cccttcggat 660
aggtcggctg ttctagactc gtccatatcc aaagagttgg aaaaaaatag caagggcccg 720
aacagtaata tttttgacac actaaacaca ctcaaggaag aaatcaaaga acttttatgt 780
catcatgtca aatattttatt gcaaaatctt acaccggagg atgcaaattt tgtgttcaat 840
agttctgtaa agtatgtaaa gaaatcatat caatactaca tacaacatc agagatggaa 900
agtgatgaat ttaagtcctt tctcactgga gtcaacatta aaatattgga gaagataatc 960
tcgtccgata ataattgttg tactccttac aaacacatca ctaatcccag gaacattatt 1020
tcgtctttac aaaaggtacg tgaactaaa cctgtttcaa aggattatcc gttcagagtg 1080
gatacggcca gagatattgt actacttcca gagacgggtg gcattttctga tctccctatc 1140
aagcccgtaa cattattgca gttggtgtcc tacatcaacg ctctcttttc cctcgagcgc 1200
cgaaatgttt tcacgcgacg cttttttaat gcagcgtgcg tcctaatttc ccagtgccta 1260
acgaatgcga acctattatc taacgacttt cctaacccca ttgaactggc agctaattgt 1320
actcgccata atcttctgag tatgaaaatg cttcaagaag gttcatccag tgaaaagaag 1380
agtaaaaaaga aggagaaaaa gaaggataaa aagaagggcg gtggtggtg tgacgattct 1440
gattcagaaa cagattcttc ctcatcatca tcatcttctt ctctctcatc ctccctctcc 1500
tcatccgaag acgaagaaga agaaaaagga gaagcagtag aaaaaggcaa gaaaactaaa 1560
cgcaaaacaa aaaagaaacc atcaaaggac gacgatttag atacaattag taaactgatt 1620
ctaaaaacag gaggttactt ccacgacacg agtgaactcg gcaataaaat tagaaattta 1680
atagacaagg atgattttgc gggcgtagcc caatatgcag taacaatcac tgagatgcaa 1740
tctacgccaa tgaatcaaa agtagtatct agtcttttag atttgataat gagactaaaa 1800
gaacaagtaa aatatagtgt tgatactgaa agtacttctt ccaccgcaa atctaataat 1860
gctttagata gtgctaaatt gacatctcaa caagtgggta caatgatggt ggattctgga 1920
gctgaattgg caaggctagc tgccttcttt ttcgtggtgg tggataatac tgtagttaat 1980
cgccatgaag cattcattct aacatcaaaa cttttaccct cgaatgaaa tagaggactt 2040
aaaactgttg tagagtcatt ctttaaaaat ttaacaatta gtaacaaagt ctctacctcg 2100
aatgaagaaa tgatgtcggg gatgccgttc gaagacgaac aacaacaaca acaatgccct 2160
caacatgaac aacaaccgga tttgaaaaga gtggtgggag aagtatttct agaaatggga 2220
aaatcaatag tgaactcatt cctttccaat aagagtgtac aattaacggc tgacgcgttc 2280
aagcaaaact actgcctat ggggaagacg ataaatttgg cggccaagat aaaacggct 2340
atatctatcg gatctaatat ctgcaccaat atactatttt ctaacctccc agaatctgta 2400
gggaataata ctgtaactgg tctaaggcta accaacctat tgaaaaacat atcccaagg 2460
gcccaagcta ataataat caaaaatgcc aacacgctcg taaacaatac gatggaccaa 2520
caaaactcag cagccatgtc catactactc tttcctccaa catcaaaaaga aacttctatc 2580
tttcccgtaa acgatccttc atctataaaa ttacaggata tgactacaat gtcaaactcg 2640
gcacgaggat tttattccat cgcagaagga tgtatcggg tcgtacgctc gaggggaattt 2700
gatgaaggag gagttaaagc gtacacttta ctagtggact caaatactat ggacatggca 2760
gttaattttg ccgctcagtc tctggaaaaa tcaatgtctg aagccttaac aaataatgag 2820
aatatgaacc cttctaattg attagaagga ggatcttttg tagacggtgc tctttcttac 2880
atgtttgaaa aaaatggcag tgactgtgaa cccacccctc tagcaaaaata tacaatgaaa 2940
gatgtatcaa atcgatactt gaaaaaattc aacaatgata aaaatacaca agatttatat 3000
aaaaatagag cagaaagagc actcgtcgaa caggttacaa ataagcctac aagtgttgtg 3060
cattcccagc tagcaaacgc catgggggta gctgttatag gcgcagcatc aatcaagctc 3120
atggaagcga aagcagcgga atctgaaatg agagcagcaa attatcaagc aacatcaaaa 3180
tctacaaatg ctattaatat aacaaataca attggaatga tacgtaatac tactcacctg 3240
tgtacgacca tcgccgtgag tgcagccgcc gacatgtcaa agctcgccaa caaccatttt 3300
atgagtgtat taacactgc aaataacagt catagcagca gacgaggaga cagatcgctc 3360
ctgatgttgc agcaacaaca accaacacat tcagcaattt tggaacagac aagaggaaga 3420
ggaggaggag tattaggatc aggaactgaa atcatgtgga atactactac acgtatgaag 3480
agagattgga tattaacat gatattctcca gaagacaaaa atactactac tactacacc 3540
agtaatgccg gccggacatt aggatattgga tctaataata ctggcataaa cacgataaaa 3600
caagacgaca agagtatgat ggataaaactt tctgagatgt ctagcttcag aacttaa 3657

```

<210> 37  
<211> 1210  
<212> PRT  
<213> SHRIMP

WO 01/38351

115

PCT/US00/28888

<400> 37

Met	Thr	Met	Asp	Asn	Val	Val	Gln	Asn	Asn	Asp	Val	Thr	Lys	Pro	Thr
1				5					10					15	
Pro	Asp	Val	Ala	Thr	Val	Thr	Thr	Ala	Thr	Glu	Lys	Arg	Gln	Ser	Cys
			20					25					30		
Lys	Glu	Lys	Lys	Asp	Gln	Leu	Lys	Ala	Glu	Cys	Pro	Gln	Val	Leu	Arg
		35				40						45			
Ala	Leu	Lys	Leu	Ser	Asn	Thr	Leu	Lys	Ala	Asn	Phe	Gly	Lys	Ser	Met
		50				55					60				
Ser	Ala	Ile	Phe	Ala	Gln	His	Leu	Val	Asp	Met	Thr	Asn	Ala	Lys	His
65					70					75					80
Phe	Lys	Asp	Pro	Lys	Thr	Lys	Lys	Ile	Leu	Asp	Gly	Ser	Ser	Ser	Ser
				85					90					95	
Asp	Ser	Glu	Glu	Glu	Glu	Thr	Ser	Ser	Ser	Ser	Lys	Arg	Lys	Arg	
			100				105						110		
Gly	Ser	Gly	Ala	Arg	Ser	Ala	Ser	Ser	Lys	Lys	Glu	Lys	Cys	Pro	Asn
		115				120						125			
Thr	Ile	Lys	Asn	Trp	Leu	Asn	Asp	Ala	Gln	Gly	Val	Phe	Arg	Gln	Phe
	130					135					140				
Ala	Asp	Ile	Ile	Ile	Asn	Leu	Pro	Ser	Phe	Asp	Asp	Leu	Arg	Asp	Glu
145					150					155					160
Val	Lys	Asp	Glu	Gln	Thr	Glu	Leu	Lys	Thr	Ile	Tyr	Asp	Leu	Tyr	Arg
				165					170					175	
Gln	Asp	Met	Glu	Lys	Val	Val	Glu	Glu	Val	Leu	Gly	Arg	Gln	Asp	Leu
			180				185						190		
Phe	Asp	His	Lys	Ser	Glu	Ile	Ala	Lys	Gly	Leu	Ala	Arg	Phe	Asp	Thr
		195					200					205			
His	Val	Ser	Leu	Leu	Pro	Ser	Asp	Arg	Ser	Ala	Val	Leu	Asp	Ser	Ser
	210				215						220				
Ile	Ser	Lys	Glu	Leu	Glu	Lys	Asn	Ser	Lys	Gly	Pro	Asn	Ser	Asn	Ile
225					230					235					240
Phe	Asp	Thr	Leu	Asn	Thr	Leu	Lys	Glu	Glu	Ile	Lys	Glu	Leu	Leu	Cys
				245					250					255	
His	His	Val	Lys	Tyr	Leu	Leu	Gln	Asn	Leu	Thr	Pro	Glu	Asp	Ala	Asn
			260				265						270		
Phe	Val	Phe	Asn	Ser	Ser	Val	Lys	Tyr	Val	Lys	Lys	Ser	Tyr	Gln	Tyr
		275					280					285			
Tyr	Ile	Gln	Thr	Ser	Glu	Met	Glu	Ser	Asp	Glu	Phe	Lys	Ser	Leu	Leu
	290					295				300					
Thr	Gly	Val	Asn	Ile	Lys	Ile	Lys	Ile	Ile	Ser	Asp	Asn	Asn	Val	
305					310					315				320	
Ala	Thr	Pro	Tyr	Lys	His	Ile	Thr	Asn	Pro	Arg	Asn	Ile	Ile	Ser	Ser
				325					330					335	
Leu	Gln	Lys	Val	Arg	Glu	Thr	Lys	Pro	Val	Ser	Lys	Asp	Tyr	Pro	Phe
			340					345					350		
Arg	Val	Asp	Thr	Ala	Arg	Asp	Ile	Val	Leu	Leu	Pro	Glu	Thr	Gly	Gly
		355					360					365			
Ile	Ser	Asp	Leu	Pro	Ile	Lys	Pro	Val	Thr	Leu	Leu	Gln	Leu	Val	Ser
	370					375					380				
Tyr	Ile	Asn	Ala	Leu	Phe	Ser	Leu	Glu	Arg	Arg	Asn	Val	Phe	Thr	Asp
385					390					395					400
Gly	Phe	Phe	Asn	Ala	Ala	Cys	Val	Leu	Ile	Ser	Gln	Cys	Leu	Thr	Asn
			405						410					415	
Ala	Asn	Leu	Leu	Ser	Asn	Asp	Phe	Pro	Lys	Pro	Ile	Glu	Leu	Ala	Ala
			420					425					430		
Asn	Val	Thr	Arg	His	Asn	Leu	Leu	Ser	Met	Lys	Met	Leu	Gln	Glu	Gly
		435					440						445		
Ser	Ser	Ser	Glu	Lys	Lys	Ser	Lys	Lys	Lys	Glu	Lys	Lys	Lys	Asp	Lys
	450					455					460				
Lys	Lys	Gly	Gly	Gly	Gly	Gly	Asp	Asp	Ser	Asp	Ser	Glu	Thr	Asp	Ser

465	470								475								480	
Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser			
				485					490					495				
Glu	Asp	Glu	Glu	Glu	Glu	Lys	Gly	Glu	Ala	Val	Glu	Lys	Gly	Lys	Lys			
			500					505					510					
Thr	Lys	Arg	Lys	Thr	Lys	Lys	Lys	Pro	Ser	Lys	Asp	Asp	Asp	Leu	Asp			
		515					520					525						
Thr	Ile	Ser	Lys	Leu	Ile	Leu	Lys	Thr	Gly	Gly	Tyr	Phe	His	Asp	Thr			
	530					535				540								
Ser	Glu	Leu	Gly	Asn	Lys	Ile	Arg	Asn	Leu	Ile	Asp	Lys	Asp	Asp	Phe			
545					550				555						560			
Ala	Gly	Val	Ala	Gln	Tyr	Ala	Val	Thr	Ile	Thr	Glu	Met	Gln	Ser	Thr			
				565					570					575				
Pro	Met	Asn	Gln	Arg	Leu	Val	Ser	Ser	Leu	Leu	Asp	Leu	Ile	Met	Arg			
			580					585					590					
Leu	Lys	Glu	Gln	Val	Lys	Tyr	Ser	Val	Asp	Thr	Glu	Ser	Thr	Ser	Ser			
		595					600					605						
Thr	Ala	Lys	Ser	Asn	Asn	Ala	Leu	Asp	Ser	Ala	Lys	Leu	Thr	Ser	Gln			
	610					615					620							
Gln	Val	Val	Thr	Met	Met	Val	Asp	Ser	Gly	Ala	Glu	Leu	Ala	Arg	Leu			
625					630					635					640			
Ala	Ala	Phe	Phe	Phe	Val	Val	Val	Asp	Asn	Thr	Val	Val	Asn	Arg	His			
				645					650					655				
Glu	Ala	Phe	Ile	Leu	Thr	Ser	Lys	Leu	Leu	Pro	Ser	Asn	Glu	Asn	Arg			
			660					665					670					
Gly	Leu	Lys	Thr	Val	Val	Glu	Ser	Phe	Phe	Lys	Asn	Leu	Thr	Ile	Ser			
		675					680					685						
Asn	Lys	Val	Ser	Thr	Ser	Asn	Glu	Glu	Met	Met	Ser	Val	Met	Pro	Phe			
	690					695					700							
Glu	Asp	Glu	Gln	Gln	Gln	Gln	Gln	Cys	Pro	Gln	His	Glu	Gln	Gln	Pro			
705					710					715					720			
Asp	Leu	Lys	Arg	Val	Val	Gly	Glu	Val	Phe	Leu	Glu	Met	Gly	Lys	Ser			
			725						730					735				
Ile	Val	Asn	Ser	Phe	Pro	Ser	Asn	Lys	Ser	Val	Gln	Leu	Thr	Ala	Asp			
		740						745					750					
Ala	Phe	Lys	Gln	Asn	Tyr	Ser	Pro	Met	Gly	Arg	Arg	Ile	Asn	Leu	Ala			
		755					760					765						
Ala	Lys	Ile	Lys	Thr	Ala	Ile	Ser	Ile	Gly	Ser	Asn	Ile	Ser	Pro	Asn			
	770					775					780							
Ile	Leu	Phe	Ser	Asn	Leu	Pro	Glu	Ser	Val	Gly	Asn	Asn	Thr	Val	Thr			
785					790					795					800			
Gly	Leu	Arg	Leu	Thr	Asn	Leu	Leu	Lys	Asn	Ile	Ser	Gln	Ser	Ala	Gln			
			805						810					815				
Ala	Asn	Asn	Ile	Ile	Lys	Asn	Ala	Asn	Thr	Leu	Val	Asn	Asn	Thr	Met			
			820					825										



WO 01/38351

PCT/US00/28888

117

Cys Glu Pro Thr Pro Leu Ala Lys Tyr Thr Met Lys Asp Val Ser Asn  
 965 970 975  
 Arg Tyr Leu Lys Lys Phe Asn Asn Asp Lys Asn Thr Gln Asp Leu Tyr  
 980 985 990  
 Lys Asn Arg Ala Glu Arg Ala Leu Val Glu Gln Val Thr Asn Lys Pro  
 995 1000 1005  
 Thr Ser Val Val His Ser Gln Leu Ala Asn Ala Met Gly Val Ala Val  
 1010 1015 1020  
 Ile Gly Ala Ala Ser Ile Lys Leu Met Glu Ala Glu Ala Ala Glu Ser  
 1025 1030 1035 1040  
 Glu Met Arg Ala Ala Asn Tyr Gln Ala Thr Ser Lys Ser Thr Asn Ala  
 1045 1050 1055  
 Ile Asn Ile Thr Asn Thr Ile Gly Met Ile Arg Asn Thr Thr His Leu  
 1060 1065 1070  
 Cys Thr Thr Ile Ala Val Ser Ala Ala Asp Met Ser Lys Leu Ala  
 1075 1080 1085  
 Asn Asn His Phe Met Ser Val Leu Asn Thr Ala Asn Asn Ser His Ser  
 1090 1095 1100  
 Ser Arg Arg Gly Asp Arg Ser Ser Leu Met Leu Gln Gln Gln Gln Pro  
 1105 1110 1115 1120  
 Thr His Ser Ala Phe Leu Glu Gln Thr Arg Gly Arg Gly Gly Gly Val  
 1125 1130 1135  
 Leu Gly Ser Gly Thr Glu Gln Thr Lys Asp His Val Glu Arg Met Lys  
 1140 1145 1150  
 Arg Asp Trp Ile Leu Asn Met Ile Ser Pro Glu Asp Lys Asn Thr Thr  
 1155 1160 1165  
 Thr Thr Thr Pro Ser Asn Ala Gly Arg Thr Leu Gly Tyr Gly Ser Asn  
 1170 1175 1180  
 Ile Thr Gly Ile Asn Thr Ile Lys Gln Asp Asp Lys Ser Met Met Asp  
 1185 1190 1195 1200  
 Lys Leu Ser Glu Met Ser Ser Phe Arg Thr  
 1205 1210

<210> 38  
 <211> 2376  
 <212> DNA  
 <213> SHRIMP

<400> 38  
 atgtctcttg ttgagaataa taccgaagaa gaaatgattc tggaaactac tgttgaaggt 60  
 gttgtgagag gagcagaagt tgcccctaga ggtgtcaaga gaccccttcc ttcttctca 120  
 tcttcttctt cagcttcgga tagtgaagat gatgaaggag gagagcaacc acaaacgaaa 180  
 ccccaaaaga agaaacgtaa tatcaatagc ggaaagtact ggaaaatcga aacaatcgaa 240  
 ccagcatctc cagaaaatgct cagtgtctgta aatgatattg ataacgtgtc caaaactatt 300  
 cctctaattc ataatagttt tggcggttcag tttaaaaaga gtgtatctga agaacagatc 360  
 aaaacgctac tcaactgaaac tattgccgtg gaatatggaa caatcacaaa tgttaaatat 420  
 tcaaccttta atcaactcga aaggaccgga gagcctctta agaagaagcg ctcaaacaat 480  
 ggaaataaca actacagata ctggcaaatt cgaattgaag ccgctgcagc agagaatgtt 540  
 acacaagctg ttttgagcgc aatcgttgaa ggaaatgaca ccgtgatcaa ggcaatcctc 600  
 ctccctgaag gggagggaat cgggcttcaa ttaacaaga gtgttagctc tcaacaagct 660  
 aaaaatattg tccaggctgc cgaatttgaa tttggacaag ttgcgcacat gaagtgtat 720  
 ttgttccaca agatggagaa ggccgatgaa tcttctaatt cttctggtga atcaccaaag 780  
 gttaagaaag taaggaggaa caagtcccag cctacaaatt cttactacac tttcactatg 840  
 atcggggatt ctcttcaaga gcgtattgat aacgcaatca aagttatcga aatgtctcca 900  
 gttaagaggc catttctcaa ttctgtctgt gctgtgaag aagacaccac tactacgacc 960  
 actacttcca ctgggtgtgt taatccacga gggattaagg atattcactt tttcgattca 1020  
 tccatttcta agggatgttt cactgtgagg aatattgttg cagcaaatgg tgaagttcca 1080  
 caagaagaat tcgtttctga gttgtacact aaccttctaa aggttgaaga gaaggtggat 1140  
 catcctactt tcaagaagct tattcatgac cgtacaatga atcgacacat taaggcatgg 1200  
 tactgcatct gccctacta caccacagga ggcgtccccc ctgcagctga taaggtttcc 1260

gctaagggaa	ttgctacct	cagaatatac	gaggacagga	ctggagtgtt	ccaattcgat	1320
ggggcacata	catctactac	cccagcacag	gcagccgagg	caactggtgc	tattcacaag	1380
tctatgctct	tccagagtc	tggaactgat	attcaaaagt	tctctgatgc	taagaaggca	1440
gaagggtttg	agcctatttc	atcgggcgaa	attgtgtacc	gttcaaagt	gagccccaat	1500
gatagcagg	caacacgctg	cttcaaattt	tattctctct	cagacgagaa	gatgaacatt	1560
gctgatgttt	tgctctatgt	tctatactat	ggacttttca	gtagtgtaca	ctttagaaga	1620
gatactatgg	aatatggtgt	tgcaaaagac	aagagtaaaa	ttatcccca	aactatcaag	1680
attaagaagg	gaggagatac	tttccatagc	gaggaggata	ttgaagtgcc	tgtaaaattc	1740
actgcaatca	cttcggagga	acttaatagg	gaatgtaaca	ccaagggaa	gaacagcctc	1800
cgagcacaca	agaaacgaaa	gagcaactct	tctactacta	ctacttcac	cacttccact	1860
tcaacaacag	caaatactcc	aaagaagacg	aagaagagcg	cttctgtgtc	cagtgcacca	1920
tttgcaaaag	ttactttaga	ttatgttgat	agtacatcat	ttgtatttta	caatatcagt	1980
aaggaaatgg	tgcagtagat	tttggctcag	gagagagtaa	agacgctaaa	ggcagtcagg	2040
aacgaagaga	agatggaaat	tgtagaggga	gaagaagcac	aagaaactta	tagagggaatc	2100
gtaaagatca	agacgaatgc	aaaggcatac	aatcttgcca	acaaaacatc	tggagtactt	2160
ttccctgctg	ataaggtgtg	tctaaagcac	acattggaag	atttggggga	tgtacttgat	2220
tttgatgttg	taaggaggga	taatgttaat	aagacggtcg	cttctactac	tactacatca	2280
tctgaaaata	aggcaagcgg	aggagatgat	gaagaaaactc	caatggaatt	tgaaactgac	2340
ggagagaagc	tggttcacga	attgttgaat	qaataa			2376

```
<210> 39
<211> 783
<212> PRT
<213> SHRIMP
```

<400> 39

Met 1	Ser	Leu	Val	Glu 5	Asn	Asn	Thr	Gln	Glu 10	Glu	Met	Ile	Thr	Thr 15	Val
Glu	Gly	Val	Val 20	Glu	Gly	Ala	Glu	Val 25	Ala	Pro	Arg	Gly	Val 30	Lys	Arg
Pro	Leu	Pro 35	Ser	Ser	Ser	Ser	Ser 40	Ser	Ser	Ala	Ser	Asp 45	Ser	Glu	Asp
Asp	Glu	Gly	Gly	Glu	Gln	Pro 55	Gln	Thr	Lys	Pro	Pro 60	Lys	Lys	Lys	Arg
Asn 65	Ile	Asn	Ser	Gly	Lys 70	Tyr	Trp	Lys	Ile 75	Glu	Thr	Ile	Glu	Pro	Asp 80
Glu	Met	Leu	Ser	Ala 85	Val	Asn	Asp	Ile	Asp 90	Asn	Val	Ser	Lys	Thr 95	Ile
Pro	Leu	Ile	Asp 100	Asn	Ser	Phe	Gly	Val 105	Gln	Phe	Lys	Lys	Ser 110	Val	Ser
Glu	Glu	Gln 115	Ile	Lys	Thr	Leu	Leu 120	Thr	Glu	Thr	Ile	Ala 125	Val	Glu	Tyr
Gly	Thr	Ile	Thr	Asn	Val	Lys 135	Tyr	Ser	Thr	Phe	Asn 140	Gln	Leu	Glu	Arg
Thr 145	Gly	Glu	Pro	Leu	Lys 150	Lys	Lys	Arg	Ser	Asn 155	Asn	Gly	Asn	Asn	Asn 160
Tyr	Arg	Tyr	Trp	Gln 165	Ile	Arg	Ile	Glu	Ala 170	Ala	Ala	Ala	Glu	Asn 175	Val
Thr	Gln	Ala	Val 180	Leu	Asp	Ala	Ile	Val 185	Glu	Gly	Asn	Asp 190	Thr	Val	Ile
Lys	Ala	Ile 195	Leu	Leu	Pro	Glu	Gly 200	Glu	Gly	Ile	Gly	Leu 205	Gln	Phe	Asn
Lys	Ser	Val	Ser	Ser	Gln	Gln 215	Ala	Lys	Asn	Ile	Val 220	Gln	Ala	Ala	Asp
Ile 225	Glu	Phe	Gly	Gln	Val 230	Ala	His	Met	Lys	Cys 235	Asn	Leu	Phe	His	Lys 240
Met	Glu	Lys	Ala	Asp 245	Glu	Ser	Ser	Asn	Ser 250	Ser	Gly	Glu	Ser	Pro 255	Lys
Val	Lys	Lys	Val 260	Arg	Arg	Asn	Lys	Ser 265	Gln	Pro	Thr	Asn	Ser 270	Tyr	Tyr
Thr	Phe	Thr	Met	Ile	Gly	Asp	Ser	Leu	Gln	Glu	Arg	Ile	Asp	Asn	Ala

Ile	Lys	Val	Ile	Glu	Met	Ser	Pro	Val	Lys	Arg	Pro	Phe	Ser	Asn	Ser
290	295	275					280				285				
Ala	Ala	Ala	Ala	Glu	Glu	Asp	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Ser	Thr
305				310						315					320
Gly	Val	Val	Asn	Pro	Arg	Gly	Ile	Lys	Asp	Ile	His	Phe	Phe	Asp	Ser
			325						330					335	
Ser	Ile	Ser	Lys	Gly	Cys	Phe	Thr	Val	Arg	Asn	Ile	Val	Ala	Ala	Asn
			340					345				350			
Gly	Glu	Val	Pro	Gln	Glu	Glu	Phe	Val	Ser	Glu	Leu	Tyr	Thr	Asn	Leu
		355					360					365			
Leu	Lys	Val	Glu	Glu	Lys	Val	Asp	His	Pro	Thr	Phe	Lys	Lys	Leu	Ile
	370					375					380				
His	Asp	Arg	Thr	Met	Asn	Arg	His	Ile	Lys	Ala	Trp	Tyr	Cys	Ile	Cys
385				390						395					400
Pro	Tyr	Tyr	Thr	Thr	Gly	Gly	Val	Pro	Pro	Ala	Ala	Asp	Lys	Val	Ser
			405						410					415	
Ala	Lys	Gly	Ile	Ala	Tyr	Ile	Tyr	Glu	Asp	Arg	Thr	Gly	Val	Phe	Gln
			420					425					430		
Phe	Asp	Gly	Ala	His	Thr	Ser	Thr	Thr	Pro	Ala	Gln	Ala	Ala	Glu	Ala
		435					440					445			
Thr	Gly	Ala	Ile	His	Lys	Ser	Met	Leu	Phe	Gln	Ser	Pro	Gly	Thr	Asp
	450					455					460				
Ile	Gln	Lys	Phe	Leu	Asp	Ala	Lys	Lys	Ala	Glu	Gly	Leu	Glu	Pro	Ile
465				470						475					480
Ser	Ser	Gly	Glu	Ile	Val	Tyr	Arg	Ser	Lys	Trp	Ser	Pro	Asn	Asp	Ser
			485						490					495	
Arg	Ala	Thr	Arg	Cys	Phe	Lys	Phe	Tyr	Ser	Ser	Ser	Asp	Glu	Lys	Met
			500					505				510			
Asn	Ile	Ala	Asp	Val	Leu	Ser	Ile	Val	His	Thr	Asp	Gly	Leu	Phe	Ser
		515					520					525			
Ser	Val	His	Phe	Arg	Lys	Asp	Thr	Met	Glu	Tyr	Gly	Val	Ala	Lys	Ser
	530					535					540				
Lys	Ser	Lys	Ile	Ile	Pro	Lys	Thr	Ile	Lys	Ile	Lys	Lys	Gly	Gly	Asp
545				550						555					560
Thr	Phe	His	Ser	Glu	Glu	Asp	Ile	Glu	Val	Pro	Val	Lys	Phe	Thr	Ala
			565						570					575	
Ile	Thr	Ser	Glu	Glu	Leu	Asn	Arg	Glu	Cys	Asn	Thr	Lys	Gly	Met	Asn
			580					585				590			
Ser	Leu	Arg	Ala	His	Lys	Lys	Arg	Lys	Ser	Asn	Ser	Ser	Thr	Thr	Thr
		595					600					605			
Thr	Ser	Thr	Thr	Ser	Thr	Ser	Thr	Thr	Ala	Asn	Thr	Pro	Lys	Lys	Thr
	610					615					620				
Lys	Lys	Ser	Ala	Ser	Ala	Ala	Ser	Asp	Pro	Phe	Ala	Lys	Leu	Thr	Leu
625				630						635					640
Asp	Tyr	Val	Asp	Ser											

WO 01/38351

120

PCT/US00/28888

Phe Glu Thr Asp Gly Glu Lys Leu Leu His Glu Leu Leu Asn Glu  
770 775 780

<210> 40  
<211> 1878  
<212> DNA  
<213> SHRIMP

<400> 40  
atgagctggt cttcttcttc atcgtgctca agttcatctg aagaggagaa tgaggtagga 60  
gtagaaggag gaggaggcag aatcggacca acagaggcca agaaaaagat cctccgtaaa 120  
cgaaagagat cttctgttaa aagtacatca tcttcttctt cctcatcatc atcatcagac 180  
gattctgact ctgatagaga agaaaaagaa ggaagaaaaac tatacgtgga tattgcagat 240  
acaaggaaac cgcccaaagt aagaaaaactg gatactcctt cacaaaacttt agagaacgac 300  
ctctacatgt ctagttoatc ttcttcttcc tcctcatctt ccgactcatc atcgtcttct 360  
ggtgaagaag aaagtgcaga tgatgatgat gacgattatg acccagataa tgtccatggt 420  
cttggttgta agaaggaaaa atctccccaa gacatagaag ctgaaaagga aaaggaggaa 480  
gagtatgaag aagaattcaa gagaatggca ttaccttcac ggataaacac atctgtagat 540  
gattgtgta tacctgatcg gattttaacc ctcttttcta cacttttgaa gaaaaatagt 600  
ttccagttct ctacgccagt ttctttcctc cggttggtga tgaagcaagt aaacgaggcc 660  
atgaattcag cattctcttc catgttatcc agttctggta tgaggttggt ggaggactct 720  
ttgggtgaca catctaaaat ttctcctttt ataacacctc aaacggatac tagtaattca 780  
tcttcatcat cgacgttcgt caataattgt acagatgaag atatcaagaa gcgtaacatt 840  
gcaatgggaa gagtagcaga attgctgtca aatattgcag ccagttctaa tgaggagaat 900  
aatttccgcc ctgtagtttc cttaatgcga ggaccaactt gtggaggttc taatgcatcc 960  
aacaagaaac tcaatagtaa caggcaaact attccccaa gttataaataa ggttatattt 1020  
tttagagaaa tacacagtgt tatcgcgcta tttttatcgt ccgtctgtgt gcagcgagca 1080  
atgaataatg acaacacaaa ttctagcgga tacgctgaag ggatggttac taaaatcttg 1140  
aatattattg gtaaaattcc ttataatgaa atgagtagag aaaaattcat atccgttgga 1200  
agagatgcac tatatttgta ccagaatgtg atcacggata tgactggccc caaacataac 1260  
aagagactcc gtatccctca acaacaagct gatttttggt acattatagc catgttggtt 1320  
aatgatgttc ccattacctc agatttactt ttaactggaa aggcgacaaa tttagtacaa 1380  
tttgcttctg ctatggtgga tcctgcatat cgcttggtg tccataaaat ggcgtctggt 1440  
ttcaatagta gttattccgt atataaagtc ctagatcttg accataaaat gttattaagg 1500  
gctaatttaa tcctatctat tttatcagct agaaataagt gtctcagtga aagaaaacct 1560  
agaacattaa ctcagagcgt gtatttggtc ctaaatcatc ttttgcgga caaattgagg 1620  
tctagtgggt tgaccagtga agagagttct ctaggaacag ccgtaaaatt ggtgtcacag 1680  
caattaatgt atgagggtgt gactcgtaaa accatcgagg acggatgtag catgattagc 1740  
ggaaactttg aggatgaaga cgggtgaaca ctgaaatgtt tgggagccga tgtaaggac 1800  
gtgaagactg ttggactatc tgcattgcta tctgatagga ttagaaaaaa cattagaaga 1860  
aacgtccctt tctactga 1878

<210> 41  
<211> 621  
<212> PRT  
<213> SHRIMP

<400> 41  
Met Ser Cys Ser Ser Ser Ser Cys Ser Ser Ser Ser Glu Glu Glu  
1 5 10 15  
Asn Glu Val Gly Val Glu Gly Gly Gly Arg Ile Gly Pro Thr Glu  
20 25 30  
Ala Lys Lys Lys Ile Leu Arg Lys Arg Lys Arg Ser Ser Val Lys Ser  
35 40 45  
Thr Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Asp Asp Ser Asp Ser  
50 55 60  
Asp Arg Glu Glu Lys Glu Gly Arg Lys Leu Tyr Val Asp Ile Ala Asp  
65 70 75 80  
Thr Arg Lys Pro Pro Lys Val Arg Lys Leu Asp Thr Pro Ser Gln Thr  
85 90 95

Leu	Glu	Asn	Asp	Leu	Tyr	Met	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser
Ser	Ser	Asp	Ser	Ser	Ser	Ser	Ser	Gly	Glu	Glu	Glu	Ser	Asp	Asp	Asp
Asp	Asp	Asp	Asp	Tyr	Asp	Pro	Asp	Asn	Val	His	Val	Leu	Gly	Cys	Lys
Lys	Glu	Lys	Ser	Pro	Gln	Asp	Ile	Glu	Ala	Glu	Lys	Glu	Lys	Glu	Glu
Glu	Tyr	Glu	Glu	Glu	Phe	Lys	Arg	Met	Ala	Leu	Pro	Ser	Arg	Ile	Asn
Thr	Ser	Val	Asp	Asp	Cys	Val	Ile	Pro	Asp	Arg	Ile	Leu	Thr	Leu	Phe
Ser	Thr	Leu	Leu	Lys	Lys	Asn	Ser	Phe	Gln	Phe	Ser	Gln	Pro	Val	Ser
Phe	Leu	Arg	Leu	Val	Met	Lys	Gln	Val	Asn	Glu	Ala	Met	Asn	Ser	Ala
Phe	Ser	Ser	Met	Leu	Ser	Ser	Ser	Gly	Met	Arg	Leu	Val	Glu	Asp	Ser
Leu	Gly	Asp	Thr	Ser	Lys	Ile	Ser	Ser	Phe	Ile	Thr	Pro	Gln	Thr	Asp
Thr	Ser	Asn	Ser	Ser	Ser	Ser	Ser	Thr	Phe	Val	Asn	Asn	Cys	Thr	Asp
Glu	Asp	Ile	Lys	Lys	Arg	Asn	Ile	Ala	Met	Gly	Arg	Val	Ala	Glu	Leu
Leu	Ser	Asn	Ile	Ala	Ala	Ser	Ser	Asn	Glu	Glu	Asn	Asn	Phe	Arg	Pro
Val	Val	Ser	Leu	Met	Arg	Gly	Pro	Thr	Cys	Gly	Gly	Ser	Asn	Asn	Lys
Lys	Leu	Asn	Ser	Asn	Arg	Gln	Thr	Ile	Pro	Gln	Val	Leu	Asn	Lys	Val
Ile	Phe	Phe	Arg	Glu	Ile	His	Ser	Val	Ile	Ala	Leu	Tyr	Leu	Ser	Ser
Val	Cys	Val	Gln	Arg	Ala	Met	Asn	Asn	Asp	Asn	Thr	Asn	Ser	Ser	Gly
Tyr	Ala	Glu	Gly	Met	Val	Thr	Lys	Ile	Leu	Asn	Ile	Ile	Gly	Lys	Ile
Pro	Tyr	Asn	Glu	Met	Ser	Arg	Glu	Lys	Phe	Ile	Ser	Val	Gly	Arg	Asp
Ala	Leu	Tyr	Leu	Tyr	Gln	Asn	Val	Ile	Thr	Asp	Met	Thr	Gly	Pro	Lys
His	Asn	Lys	Arg	Leu	Arg	Ile	Pro	Gln	Gln	Gln	Ala	Asp	Phe	Cys	Tyr
Ile	Ile	Ala	Met	Leu	Val	Asn	Asp	Val	Pro	Ile	Thr	Ser	Asp	Leu	Leu
Leu	Thr	Gly	Lys	Ala	Thr	Asn	Leu	Val	Gln	Phe	Ala	Ser	Ala	Met	Val
Asp	Pro	Ala	Tyr	Arg	Leu	Ala	Val	His	Lys	Met	Ala	Ser	Val	Phe	Asn
Ser	Ser	Tyr	Ser	Val	Tyr	Lys	Val	Leu	Asp	Leu	Asp	His	Lys	Met	Leu
Leu	Arg	Ala	Asn	Leu	Ile	Leu	Ser	Ile	Leu	Ser	Ala	Arg	Asn	Lys	Cys
Leu	Ser	Lys	Pro	Arg	Thr	Leu	Thr	Gln	Ser	Val	Tyr	Leu	Phe	Leu	Asn
His	Leu	Leu	Arg	Asn	Lys	Leu	Arg	Ser	Ser	Gly	Leu	Thr	Ser	Glu	Glu
Ser	Ser	Leu	Gly	Thr	Ala	Val	Lys	Leu	Val	Ser	Gln	Gln	Leu	Met	Tyr
Glu	Gly	Val	Thr	Arg	Gln	Thr	Ile	Glu	Asp	Gly	Cys	Ser	Met	Ile	Ser
Gly	Asn	Phe	Glu	Asp	Glu	Asp	Gly	Val	Thr	Leu	Lys	Cys	Leu	Gly	Ala

580 585 590  
 Asp Val Lys Asp Val Lys Thr Val Gly Leu Ser Ala Leu Leu Ser Asp  
 595 600 605  
 Arg Leu Arg Lys Asn Ile Arg Arg Asn Val Pro Phe Tyr  
 610 615 620

<210> 42  
 <211> 3303  
 <212> DNA  
 <213> SHRIMP

<400> 42  
 atgttttgaa gttctgctaa taactttaat ggtgacaaga aatcttcctc atcatcatca 60  
 gctgccgcac catctgacga tcagcagtta ggtccccctg gactgtctac tgctgatttc 120  
 aagaaagttg ctgccattct tgccaacaga acagagtcoc tatatctgtt gccagattct 180  
 cccaacttta agaattgtgat taacaaccct aaccagatat ctattgtgcc gtttctgggt 240  
 tcatcgaagg cagctgaaag tggtagtgca aacaagaatg aaaaccaggc tgaaaattct 300  
 tctaaaggag gaagtgatgg aaagaaatca tcgcagcaga acaagtttaa cctattgaac 360  
 aaggtagagg ctgaagaaat ggccctttaa cgtgtggctg aactcattgc ggatactcct 420  
 cctctaaaag ataattccatt gagggatgac cctgatgcta ttccatcacg caaccatgg 480  
 gtcaaattga ctcaagaaga tctggaatat cttttctggg aagcagtaac aattgaagtc 540  
 tccaatgata ggagtatccg tagtggaaga tatcttcaag ccagtgaagt gggggagaat 600  
 ccattcctaa tgaccatcag tgttgacatt agaactcctg aaagaatggc acttaatgtc 660  
 gtgtggttct tcaacagatt ctccgcgatg gtttctggac ttggagtaga aaacagggcc 720  
 aattccacct atgtagcaac tagcgtatgt attgcccaga tctgggtaga gatgtcctc 780  
 aagaacttta tttctggaga aaatgtgcc caggcactga agtatttgaa ggaacattat 840  
 gagcatgttt ataacaagat cagtaaattgt ggacgtcagc catcctattt tgttgttgaa 900  
 tttgaaagag tggacaacac gattggattt gtcaattctg atactgaaca taatggatca 960  
 tcatacatgg aatacagggtg ctttgacaca atcaggaaga atgcatcatc tgggccaagt 1020  
 ggaggcggaa agagtgggtg cctgtcttct ggaacattct tcattgataa cgaaatgggg 1080  
 aataacaata gtagtgcagc tgcagcttct gccoctgctg tttctgctgg agtttctcca 1140  
 tcactttctc catttagcag tgatggagat gatgatgacg acgattgtag tggtagtgac 1200  
 gtgtggggga agaagatgat attcaacaca tcaggagacg gatcaggaga atcttctggg 1260  
 cagaacgggtg gtggtgcac aacttacaag agatttaggt gtggagaaaa tactgcttct 1320  
 ctttctcaga aggaaaatgt tegtctcatg gcgatgccaa agggaaatga agataaaca 1380  
 ctactcaaga acattatcaa ctttctgaat agtgcactta actctgttga aaaccatgta 1440  
 atgtgtacag atgaaaatat cttcgatgag gatcaagccg agcattatac gtcaaacaag 1500  
 gaactataca aggctattgt ttgctccaac ccagccaatg tgtacagagt aatggttgaa 1560  
 ctgtttgtca acctatttct tccccgtctt aggaacccaa ttgtgagtga cattgaaact 1620  
 gtacaaaatc ttccttcaaa caatggatcg gtgagaacaa agaagatggt tgaacatgga 1680  
 tgtacagata tgcgctatga catccctcca atgcaaaagg gaaagattcg tcttagtgcc 1740  
 aagagagctt gtgagtgcag aaaactgtgc aaggacgtga ggtgctttga caagtctaga 1800  
 gaggcaaatc tcacccttag ccagaaggca ggaagggagg tggagggaacc ttttccccgc 1860  
 aaccacaatt ctacacaggag taacgctcac gactttactt tctatgacaa gtacagggca 1920  
 aggatgaaca agctcaagaa ggattcaaag aagaaggtaa agaagattga cacctttaca 1980  
 acaacagacg attttctcct gcaagatagg aacgcttttg atctacttag aaagtgtctc 2040  
 ctttctgcct ctcttcatca cattttctgt cctgatgtgc ttatggtgca tagaggagac 2100  
 agcttcaata ttaactttgc aaacaacaaa ctcgagtgtc ataataaacg taacggaatt 2160  
 gaagaagtta cttcttcaca aacggtcaac gccaaaggaa cacttgagga tattacaaaa 2220  
 attaaaatga agagagggga tgatattata gatgttgtga agagtaaagg actttctttg 2280  
 agggaattct ctaagaaggt tagtaagatt gtgagaaggt ttaatgagat cacaaacca 2340  
 ctctgcaaca actgcaacgt taactcttct aatggagatg tggatttcca cgtctttact 2400  
 tctgtgtgtg tctacatcca caacattatt cctgtgctcg aagatatctc catttttgca 2460  
 gaattgggtg aagaattgac caagcttggt aaggagtgtg gagacgtggc tggagaggac 2520  
 aagacatatg atgatattat ccgcaattac gaaattactg taaagtactt taagctcttt 2580  
 aatgcactcg ttaaatctctg tcacaggaat tataatgtgg cagtaacctc tgccattaac 2640  
 aggagagggt acatgtgcat ggtgagcaac ctgttcgggt attattgtaa gctgtctgat 2700  
 aacgctatcc agtatcacga atcactatgc tctttgcact ctagcatctc ttatgcagac 2760  
 tattatacgt ctgcgaataa caattctgaa gatggaggag gaaactcttc ttcagaaaag 2820  
 agcaatgcag atgtagccaa gactatggcc tctttctatg accagttcga taagagtga 2880

Met 1	Phe	Gly	Ser	Ser 5	Ala	Asn	Asn	Phe	Asn 10	Gly	Asp	Lys	Lys	Ser 15	Ser
Ser	Ser	Ser	Ser 20	Ala	Ala	Ala	Ser	Ser 25	Asp	Asp	Gln	Gln	Leu 30	Gly	Pro
Leu	Gly	Leu 35	Ser	Thr	Ala	Asp	Phe 40	Lys	Lys	Val	Ala	Ala 45	Ile	Leu	Ala
Asn 50	Arg	Thr	Glu	Ser	Leu	Tyr 55	Leu	Leu	Pro	Asp	Ser	Pro	Asn	Phe	Lys
Asn 65	Val	Ile	Asn	Asn	Pro	Asn 70	Gln	Ile	Ser	Ile	Val 75	Pro	Phe	Leu	Gly 80
Ser	Ser	Lys	Ala	Ala 85	Glu	Ser	Gly	Ser	Ala 90	Asn	Lys	Asn	Glu	Asn 95	Gln
Ala	Glu	Asn	Ser 100	Ser	Lys	Gly	Gly	Ser 105	Asp	Gly	Lys	Lys	Ser 110	Ser	Gln
Gln	Asn	Lys 115	Phe	Asn	Leu	Leu	Asn 120	Lys	Val	Glu	Ala	Glu 125	Met	Ala	
Phe 130	Lys	Arg	Val	Ala	Glu	Leu 135	Ile	Ala	Asp	Thr	Pro	Pro	Ser	Lys	Asp
Asn 145	Pro	Leu	Arg	Asp	Asp	Pro 150	Asp	Ala	Ile	Pro	Ser	Arg	Asn	Pro	Trp 160
Val	Lys	Leu	Thr	Gln 165	Lys	Asn	Leu	Glu	Tyr 170	Leu	Phe	Trp	Glu	Ala 175	Val
Thr	Ile	Glu	Val 180	Ser	Asn	Asp	Arg	Ser 185	Ile	Arg	Ser	Gly	Arg 190	Tyr	Leu
Gln	Ala	Ser 195	Glu	Val	Gly	Glu	Asn 200	Pro	Phe	Leu	Met	Thr 205	Ile	Ser	Val
Asp	Ile	Arg	Ile	Leu	Gln	Arg 215	Met	Ala	Leu	Asn	Val 220	Val	Trp	Phe	Phe
Asn 225	Arg	Phe	Phe	Arg	Met 230	Val	Ser	Gly	Leu	Gly	Val	Glu	Asn	Arg	Ala 240
Asn	Ser	Thr	Tyr	Val 245	Ala	Thr	Ser	Asp	Ala 250	Ile	Ala	Gln	Ile	Trp 255	Val
Glu	Met	Leu	Leu 260	Lys	Asn	Phe	Ile	Ser 265	Gly	Glu	Asn	Val	Pro 270	Gln	Ala
Leu	Lys	Tyr 275	Leu	Lys	Glu	His	Tyr 280	Glu	His	Val	Tyr	Asn 285	Lys	Ile	Ser
Lys	Cys	Gly	Arg	Gln	Pro	Ser 295	Tyr	Phe	Val	Val	Glu	Phe	Glu	Arg	Val
Asp 305	Asn	Thr	Ile	Gly	Phe 310	Val	Asn	Ser	Asp	Thr	Glu	His	Asn	Gly	Ser
Ser	Tyr	Met	Glu	Tyr 325	Arg	Cys	Phe	Asp	Thr 330	Ile	Arg	Lys	Asn	Ala	Ser
Ser	Gly	Pro	Ser 340	Gly	Gly	Gly	Lys	Ser 345	Gly	Val	Leu	Ser	Ser 350	Gly	Thr
Phe	Phe	Ile 355	Asp	Asn	Glu	Met	Gly 360	Asn	Asn	Asn	Ser	Ser	Ala	Ala	Ala

WO 01/38351

124

PCT/US00/28888

Ala	Ser	Ala	Pro	Ala	Val	Ser	Ala	Gly	Val	Ser	Pro	Ser	Leu	Ser	Pro
	370					375					380				
Phe	Ser	Ser	Asp	Gly	Asp	Asp	Asp	Asp	Asp	Asp	Cys	Ser	Gly	Asp	Asp
385					390					395					400
Val	Trp	Gly	Lys	Lys	Met	Ile	Phe	Asn	Thr	Ser	Gly	Asp	Gly	Ser	Gly
				405					410					415	
Glu	Ser	Ser	Gly	Gln	Asn	Gly	Gly	Gly	Ala	Ser	Thr	Tyr	Lys	Arg	Phe
			420					425					430		
Arg	Cys	Gly	Glu	Asn	Thr	Ala	Ser	Leu	Ser	Gln	Lys	Glu	Asn	Val	Arg
		435					440					445			
Leu	Met	Ala	Met	Pro	Lys	Gly	Asn	Glu	Asp	Lys	Gln	Leu	Leu	Lys	Asn
	450					455					460				
Ile	Ile	Asn	Phe	Leu	Asn	Ser	Ala	Leu	Asn	Ser	Val	Glu	Asn	His	Val
465					470					475					480
Met	Cys	Thr	Asp	Glu	Asn	Ile	Phe	Asp	Glu	Asp	Gln	Ala	Glu	His	Tyr
				485					490					495	
Thr	Ser	Asn	Lys	Glu	Leu	Tyr	Lys	Ala	Ile	Val	Cys	Ser	Asn	Pro	Ala
			500					505					510		
Asn	Val	Tyr	Arg	Val	Met	Val	Glu	Leu	Phe	Val	Asn	Leu	Ile	Leu	Pro
	515					520						525			
Arg	Leu	Arg	Asn	Pro	Ile	Val	Ser	Asp	Ile	Glu	Thr	Val	Gln	Asn	Leu
	530					535					540				
Pro	Ser	Asn	Asn	Gly	Ser	Val	Arg	Thr	Lys	Lys	Met	Val	Glu	His	Gly
545					550					555					560
Cys	Thr	Asp	Met	Arg	Tyr	Asp	Ile	Pro	Pro	Tyr	Ala	Lys	Gly	Lys	Ile
				565					570					575	
Arg	Leu	Ser	Ala	Lys	Arg	Ala	Cys	Glu	Cys	Arg	Lys	Leu	Cys	Lys	Asp
			580					585					590		
Val	Arg	Cys	Phe	Asp	Lys	Ser	Arg	Glu	Ala	Asn	Leu	Thr	Pro	Ser	Gln
	595						600					605			
Lys	Ala	Gly	Arg	Glu	Val	Glu	Glu	Pro	Phe	Pro	Arg	Asn	His	Asn	Ser
	610					615					620				
His	Arg	Ser	Asn	Ala	His	Asp	Phe	Thr	Phe	Tyr	Asp	Lys	Tyr	Arg	Ala
625					630					635					640
Arg	Met	Asn	Lys	Leu	Lys	Lys	Asp	Ser	Lys	Lys	Lys	Val	Lys	Lys	Ile
				645					650					655	
Asp	Thr	Phe	Thr	Thr	Thr	Asp	Asp	Phe	Leu	Leu	Gln	Asp	Arg	Asn	Ala
			660					665					670		
Phe	Asp	Leu	Leu	Arg	Lys	Cys	Phe	Leu	Ser	Ala	Ser	Leu	His	His	Ile
	675						680					685			
Phe	Cys	Pro	Asp	Val	Leu	Met	Val	His	Arg	Gly	Asp	Ser	Phe	Asn	Ile
	690					695					700				
Asn	Phe	Ala	Asn	Asn	Lys	Leu	Glu	Cys	Tyr	Asn	Glu	Arg	Asn	Gly	Ile
705					710					715					720
Glu	Glu	Val	Thr	Ser	Ser	Gln	Thr	Val	Asn	Ala	Lys	Glu	Ala	Leu	Glu
				725					730					735	
Asp	Ile	Thr	Lys	Ile	Lys	Met	Lys	Arg	Gly	Asp	Asp	Ile	Ile	Asp	Val
			740					745					750		
Val	Lys	Ser	Lys	Gly	Leu	Ser	Leu	Arg	Glu	Phe	Ser	Lys	Lys	Val	Ser
	755						760					765			
Lys	Ile	Val	Arg	Arg	Phe	Asn	Glu	Ile	Thr	Asn	Gln	Leu	Cys	Asn	Asn
	770					775						780			
Cys	Asn	Val	Asn	Ser	Ser	Asn	Gly	Asp	Val	Asp	Phe	His	Val	Phe	Thr
785					790					795					800
Ser	Val	Cys	Val	Tyr	Ile	His	Asn	Ile	Ile	Pro	Val	Leu	Glu	Asp	Ile
				805					810					815	
Ser	Ile	Phe	Ala	Glu	Leu	Gly	Glu	Glu	Leu	Thr	Lys	Leu	Val	Lys	Glu
			820					825					830		
Cys	Arg	Asp	Val	Ala	Gly	Glu	Asp	Lys	Thr	Tyr	Asp	Asp	Ile	Ile	Arg
	835					840						845			
Asn	Tyr	Glu	Ile	Thr	Val	Lys	Tyr	Phe	Lys	Leu	Phe	Asn	Ala	Leu	Val



```

      850              855              860
Lys Phe Cys His Arg Asn Tyr Asn Val Ala Val Thr Ser Ala Ile Asn
865              870              875              880
Arg Arg Gly Tyr Met Cys Met Val Ser Asn Leu Val Gly Tyr Tyr Cys
      885              890              895
Lys Leu Ser Asp Asn Ala Ile Gln Tyr His Glu Ser Leu Cys Ser Leu
      900              905              910
His Ser Ser Ile Ser Tyr Ala Asp Tyr Tyr Thr Ser Arg Asn Asn Asn
      915              920              925
Ser Glu Asp Gly Gly Gly Asn Ser Ser Ser Glu Lys Ser Asn Ala Asp
      930              935              940
Val Ala Lys Thr Met Ala Ser Phe Tyr Asp Gln Phe Asp Lys Ser Glu
945              950              955              960
Asp Ser Lys Lys Asn Lys Asn Lys Thr Ser Asn Glu Ile Leu Ile Lys
      965              970              975
Met Phe Gln Met Asp Arg Val Leu Asp Gly Met Asp Asp Asp Asp
      980              985              990
Glu Asp Ser Asp Ser Ser Ser Ser Glu Asn Glu Glu Glu Glu Glu
      995              1000              1005
Glu Glu Ile Val Lys Lys Pro Ala Lys Lys Arg Lys Val Glu Asp Val
1010              1015              1020
Asp Ser Asn Lys Lys Thr Leu Pro Lys Glu Pro Ala Val Lys Lys Val
1025              1030              1035              1040
Lys Gln Glu Glu Asp Val Glu Met Glu Glu Val Lys Glu Ala Ala Ala
      1045              1050              1055
Glu Glu Glu Lys Lys Glu Glu Gln Glu Ala Lys Glu Glu Asp Ala Thr
      1060              1065              1070
Glu Tyr Asp Asp Asp Thr Glu Glu Asp Glu Lys Ala Val Ala Ser Asp
      1075              1080              1085
Glu Asp Glu Asp Asp Glu Asp Ser Lys Ala Ile Phe
1090              1095              1100

```

<210> 44  
 <211> 618  
 <212> DNA  
 <213> SHRIMP

```

<400> 44
atggcaggaa tcggtagaag agataataga cccgtattac atctggacat tgatccaaat 60
aaggaaatac cctataacgt tccaccaacc cctattatct gtgaaaaaaa ccccttcgtg 120
tttaacatgc agaagtgtct agactgtgct cctttccctc cctaccccg cactgagaag 180
cctttccctc cataccctgg tactgcagta gaagaggagg agaagcaaaa ggaaattgag 240
gagcttctgg ttgaccaatc tttccctccc ccattccctg gaaataagct gagagatatc 300
cccagaacct accctctcga atttcccgag aagaaggaga aggatttccc ttgcgttgac 360
actaccgggc acagcgatat ccccttcacg gatctggaga aaacccccacc ccgtagtgc 420
gttaggcacg gttaccacta cttaatcaac cccaacaagg ttggggagct taaccatatc 480
gttggttaagc tcaactgaaa gcaagaaaac ctgaacaaat tgggtgtgga tgttgatgac 540
gttggtgatta atctgtcaag cactttgaag gaacttgaga agctgcgagc tggcctgtgc 600
aagttctcaa aaaactag                                     618

```

<210> 45  
 <211> 205  
 <212> PRT  
 <213> SHRIMP

```

<400> 45
Met Ala Gly Ile Gly Arg Arg Asp Asn Arg Pro Val Leu His Leu Asp
  1              5              10              15
Ile Asp Pro Asn Lys Glu Ile Pro Tyr Asn Val Pro Pro Thr Pro Ile
      20              25              30

```



atgacgagac	atggtgtgct	tgttccaaaa	ggcgtttcta	ggcatgttat	cctaggaaat	60
gtcgactaca	ctttctgtac	tactgacaat	aatttgtgtca	gtctagacat	cgatttcaag	120
gacaaatatc	cagacaaaaa	cattcagtta	ttgaacaaga	aattgggtaa	gaaaaacagca	180
aagaaaataa	agaaggaaga	tgcacctgaa	acaaaaggaaa	atagtagcga	agacatatatc	240
gccaccaagg	aattcgaaaca	gacataaaaa	ggttctacaga	caaaaaaagg	tgccaccgag	300
gaaaaacgcca	tcgcggcgcg	agctgccgct	gccactgctg	ctgcggtaga	aaaggctatg	360
ctatcagaaa	gtgaaggaaa	atcaatggtc	atcaacagag	ctagaatggg	gctttctaag	420
cgagacactg	ccagaaaaa	gttcactgca	ttgaagaaca	gggaatcttt	cttcagtgtt	480
ttgatattgt	aaactggatc	agtgatagtt	gtcgggtctt	aagatccttc	gcttcaaaaa	540
ttgtgtgtga	ttaaagccac	gactgatatt	gctgatattc	tacagaaaaa	catcagtgtg	600
gctaacgtgt	ctatagtga	tacagtgtcc	acttttaata	gattccactt	gaactttatt	660
cgactcggga	aattcttcga	aagaaattgc	atctcttaca	gttataaccc	agaaacgttc	720
cccggtatgt	ttttcaagct	gcgagtcccc	gcaaagcctc	tcttgccctg	agagactata	780
ggggaatact	acacaaaggt	tgcfaatgta	cgcgatagta	aggatcccaa	ttttaaaatg	840
ctcgtactgt	tgaggataaa	aactcgatta	acttttaaag	ttgggaaaat	tactgtgctc	900
ggagaaggag	agagtgggtg	cggtgatgtt	tctgtcgtat	ccaaattact	atttggttta	960
ttccattact	ttatggacaa	caacattaaa	atgtccccc	aagaagcaca	aagagtcaga	1020
gaaaaaatag	gcaccccgca	tctagaatgg	tacttgtaca	ttgacatgtt	gctccactcc	1080
taccgctagc	tcaaaccatc	ggccgagcaa	gtgaaaaggg	cgatgggtga	ccaacaacat	1140
attcttgaag	tggataggac	atactatgga	acaaagaaca	gtatggacgc	tgccatgtct	1200
gcaaattttg	tgccttcaaa	agaagagagt	atctccttca	ttaaaaaaat	aagatcacaa	1260
caactttttg	gacattttgt	taaaccitca	aaagaaacta	ctcgacgtgc	tatagacaca	1320
ctttctttcg	atcctataaa	ccaagacagg	tgggtggaata	aaaatgacca	atattccggt	1380
aaagagagat	gtgaccggtt	tctcttgtca	cgttttagtg	ctgtttctga	aaatacaaac	1440
agtatgatga	atagtcgcat	tctctgtcaa	gggaaatggg	ggctagatga	aaacgaatac	1500
aaggataaac	ttgatcata	tgtggatttg	tgtacagaag	aaatagtgga	ggaatgtgaa	1560
tcaaagggtt	ttattgcctc	cccatttttg	aggaagcacc	agaaggaaaa	aataccaacg	1620
ccttatgttt	tattagcgag	agcctgtaat	caaaaaaatg	gtaacaaaa	gagtattaac	1680
aataatagta	actatttgtc	gggttcaagt	agggcgaa	ggaatgcaaa	actacaggaa	1740
aaacaccgtg	taactttagc	caggttgaa	acgatgatgt	cgctgtaccg	atttttgaa	1800
aattacatct	caacagacat	tgcacctgat	tttgccaagt	tatttggtaa	tgatgtatat	1860
agtttattac	atttaattgac	aaacttgctt	aaatcccgtg	gacatgctct	aacatacaac	1920
gaaaggggcc	tttcaagtaa	tgaagtaca	tataaaacac	ctggaaatgc	atactttagt	1980
actctatttg	aaaaatccat	tataaaaac	caagaaactg	ctaataaagg	taacaataga	2040
aacagtaaat	tttctcgaat	cggacagaag	aagagctctt	tctctgtcaa	cgcgtgtggt	2100
gtcaatttga	acaagggtag	tgatgaaatc	ataaagggtt	tttgtaacag	ttgcgatcaa	2160
aatagtacca	gttacataga	gaatgcatta	tctgacatta	acagagacaa	gaagattaaa	2220
cgtttttaag	cagctgcac	ccatcccgca	gtgaagcaag	aattggtaga	ttctttatcc	2280
tctcttctac	ctccttcttc	tcttcttctt	cagacgtcta	acaagaacaa	tagatgcacc	2340
cctagtgaatt	ttatagatta	tgtgtacaaa	tctactgacq	aaacaacagg	tgctccaaag	2400

WO 01/38351

128

PCT/US00/28888

gtgggcttag tgtttaaaat gtgtgatatt cttgcacccct tagcaagcag gagaggggatg 2460  
gaagatcgct ccacagccaa ctatagaacc tccttacatt cagctactca aaataaaaacc 2520  
aatttgaata aactattagt ttctgctatc aaggaaacag gagccactga aactgaagca 2580  
cagatattca acaagattat tggtagtgaa aagggaactat caattctctg tcaacttggtg 2640  
gaaaggagga acaaagacaa taatgtcttc gactgatttg tctaa 2685

<210> 49  
<211> 886  
<212> PRT  
<213> SHRIMP

<400> 49

Met	Thr	Arg	His	Gly	Val	Leu	Val	Pro	Lys	Gly	Arg	Ser	Arg	His	Val
1				5					10					15	
Ile	Leu	Gly	Asn	Val	Asp	Tyr	Thr	Phe	Cys	Thr	Thr	Asp	Asn	Asn	Cys
			20					25					30		
Val	Ser	Leu	Asp	Ile	Asp	Phe	Lys	Asp	Asn	Ile	Thr	Asp	Gln	Asn	Ile
		35					40					45			
Gln	Leu	Leu	Asn	Lys	Lys	Leu	Gly	Lys	Lys	Thr	Ala	Lys	Lys	Ile	Lys
	50					55					60				
Lys	Glu	Asp	Ala	Pro	Glu	Thr	Lys	Glu	Asn	Ser	Asp	Glu	Asp	Ile	Tyr
65					70					75					80
Ala	Thr	Lys	Glu	Phe	Glu	Gln	Thr	Ile	Lys	Gly	Leu	Gln	Thr	Lys	Lys
				85					90					95	
Gly	Ala	Thr	Glu	Asn	Ala	Ile	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Thr
			100				105							110	
Ala	Ala	Ala	Val	Glu	Lys	Ala	Met	Leu	Ser	Glu	Ser	Glu	Gly	Lys	Ser
		115					120					125			
Met	Val	Ile	Asn	Arg	Ala	Arg	Met	Val	Leu	Ser	Lys	Arg	Asp	Thr	Ser
	130					135					140				
Gln	Lys	Gln	Phe	Thr	Ala	Leu	Lys	Asn	Arg	Glu	Ser	Phe	Phe	Ser	Val
145					150					155					160
Leu	Ile	Phe	Glu	Thr	Gly	Ser	Val	Ile	Val	Val	Gly	Leu	Gln	Asp	Pro
				165					170					175	
Ser	Leu	Thr	Lys	Leu	Cys	Val	Ile	Lys	Ala	Thr	Thr	Asp	Ile	Ala	Asp
			180					185					190		
Ile	Leu	Gln	Lys	Asn	Ile	Ser	Val	Ala	Asn	Val	Ser	Ile	Val	Asn	Thr
		195					200					205			
Val	Ser	Thr	Phe	Asn	Arg	Phe	His	Leu	Asn	Phe	Ile	Arg	Leu	Gly	Lys
		210				215					220				
Phe	Phe	Glu	Arg	Asn	Cys	Ile	Ser	Tyr	Ser	Tyr	Asn	Pro	Glu	Thr	Phe
225					230					235					240
Pro	Gly	Met	Phe	Phe	Lys	Leu	Arg	Val	Pro	Ala	Lys	Pro	Leu	Leu	Pro
				245					250					255	
Gly	Glu	Thr	Ile	Gly	Glu	Tyr	Tyr	Thr	Lys	Val	Ala	Met	Met	Arg	Asp
			260					265					270		
Ser	Lys	Asp	Pro	Asn	Phe	Lys	Met	Ser	Asp	Trp	Leu	Arg	Ile	Lys	Thr
		275					280					285			
Ala	Leu	Thr	Phe	Lys	Val	Gly	Lys	Ile	Thr	Val	Leu	Gly	Glu	Gly	Glu
		290				295					300				
Ser	Gly	Cys	Gly	Asp	Val	Ser	Val	Val	Ser	Lys	Leu	Leu	Phe	Gly	Leu
305					310					315					320
Phe	His	Tyr	Phe	Met	Asp	Asn	Asn	Ile	Lys	Met	Ser	Pro	Lys	Glu	Ala
				325					330					335	
Gln	Arg	Val	Arg	Glu	Lys	Tyr	Gly	Ile	Pro	His	Leu	Glu	Trp	Tyr	Leu
			340					345					350		
Tyr	Ile	Asp	Met	Leu	Leu	His	Ser	Tyr	Pro	Tyr	Val	Lys	Pro	Ser	Ala
		355					360					365			
Glu	Gln	Val	Lys	Arg	Ala	Met	Val	Asp	Gln	Gln	His	Glu	Val	Asp	Arg
		370				375					380				
Thr	Tyr	Tyr	Gly	Thr	Lys	Asn	Ser	Met	Asp	Ala	Ala	Met	Ser	Ala	Asn

385					390					395				400	
Leu	Val	Pro	Ser	Lys	Glu	Glu	Ser	Ile	Ser	Phe	Ile	Lys	Lys	Ile	Arg
				405					410					415	
Ser	Gln	Gln	Leu	Phe	Gly	His	Leu	Cys	Lys	Pro	Ser	Lys	Glu	Thr	Thr
			420					425					430		
Arg	Arg	Ala	Ile	Asp	Thr	Leu	Ser	Phe	Asp	Pro	Ile	Asn	Gln	Asp	Arg
		435					440					445			
Trp	Trp	Asn	Lys	Asn	Asp	Gln	Tyr	Tyr	Gly	Lys	Glu	Arg	Cys	Asp	Pro
	450					455					460				
Phe	Ser	Val	Ala	Arg	Leu	Val	Ser	Val	Ser	Glu	Asn	Thr	Asn	Ser	Met
465					470					475					480
Met	Asn	Ser	Arg	Ile	Ser	Cys	Gln	Gly	Lys	Trp	Trp	Leu	Asp	Glu	Asn
			485						490					495	
Glu	Tyr	Lys	Asp	Lys	Leu	Asp	His	Ile	Val	Asp	Leu	Cys	Thr	Glu	Glu
			500					505						510	
Ile	Val	Glu	Glu	Cys	Glu	Ser	Lys	Gly	Phe	Ile	Asp	Phe	Leu	Arg	Lys
		515					520					525			
His	Gln	Lys	Glu	Lys	Ile	Pro	Thr	Pro	Tyr	Val	Leu	Leu	Ala	Arg	Ala
	530					535					540				
Cys	Asn	Gln	Lys	Asn	Gly	Asn	Lys	Met	Ser	Ile	Asn	Asn	Asn	Ser	Asn
545					550					555					560
Tyr	Leu	Ser	Gly	Ser	Ser	Arg	Ala	Lys	Arg	Asn	Ala	Lys	Leu	Gln	Glu
			565						570					575	
Lys	His	Arg	Val	Thr	Leu	Ala	Arg	Leu	Asn	Thr	Met	Met	Ala	Ser	Tyr
			580					585					590		
Arg	Phe	Leu	Asn	Asn	Tyr	Ile	Ser	Thr	Asp	Ile	Ala	Pro	Asp	Phe	Ala
	595						600					605			
Lys	Leu	Phe	Gly	Asn	Asp	Val	Tyr	Ser	Leu	Leu	His	Leu	Met	Thr	Asn
	610					615					620				
Leu	Lys	Ser	Arg	Gly	His	Ala	Leu	Thr	Tyr	Asn	Glu	Arg	Ala	Leu	Ser
625					630					635					640
Ser	Asn	Glu	Ser	Thr	Tyr	Lys	Thr	Pro	Gly	Asn	Ala	Tyr	Phe	Ser	Thr
			645						650					655	
Leu	Phe	Glu	Lys	Ser	Ile	Ile	Asn	Asn	Gln	Glu	Thr	Ala	Asn	Lys	Gly
			660					665					670		
Asn	Asn	Arg	Lys	Arg	Lys	Phe	Ser	Arg	Ile	Gly	Gln	Glu	Lys	Ser	Ser
	675						680					685			
Phe	Leu	Cys	Asn	Ala	Cys	Gly	Val	Asn	Leu	Asn	Lys	Gly	Ser	Asp	Glu
	690					695					700				
Ile	Ile	Lys	Gly	Ile	Cys	Thr	Ser	Cys	Asp	Gln	Asn	Ser	Thr	Ser	Tyr
705					710					715					720
Ile	Glu	Asn	Ala	Leu	Ser	Asp	Ile	Asn	Arg	Asp	Lys	Lys	Ile	Lys	Arg
			725						730					735	
Phe	Lys	Ala	Ala	Ala	Thr	His	Pro	Pro	Val	Lys	Gln	Glu	Leu	Val	Asp
			740					745					750		
Ser	Leu	Ser	Ser	Ser	Ser	Ser	Pro	Ser	Ser	Ser	Ser	Ser	Gln	Thr	Ser
	755						760						765		
Asn	Lys	Asn	Asn	Arg	Cys	Thr	Pro	Ser	Asp	Phe	Ile	Asp	Tyr	Val	Tyr
	770					775					780				
Lys	Phe	Thr	Asp	Glu	Thr	Thr	Gly	Ala	Pro	Lys	Val	Gly	Leu	Val	Phe
785					790					795					800
Lys	Met	Cys	Asp	Ile	Leu	Ala	Ser	Leu	Ala	Ser	Arg	Arg	Gly	Met	Glu
			805						810					815	
Asp	Arg	Pro	Thr	Ala	Asn	Tyr	Arg	Thr	Ser	Leu	His	Ser	Ala	Thr	Gln
			820					825					830		
Asn	Lys	Thr	Asn	Leu	Asn	Lys	Leu	Leu	Val	Ser	Ala	Ile	Lys	Glu	Thr
	835						840					845			
Gly	Ala	Thr	Glu	Thr	Glu	Ala	Gln	Ile	Phe	Asn	Lys	Ile	Ile	Gly	Ser
	850					855					860				
Glu	Lys	Gly	Leu	Ser	Ile	Leu	Cys	Gln	Leu	Val	Glu	Arg	Arg	Asn	Lys
865					870					875					880

WO 01/38351

130

PCT/US00/28888

Asp Asn Asn Val Phe Asp  
885

<210> 50  
<211> 1268  
<212> DNA  
<213> SHRIMP

<400> 50  
atgtcttcga ctgatttgtc taagaatgcc ttccatgact ggggtggtatc aaagacagat 60  
tgtgaggtgt ttgatgtaca ctgtgagacg gacagagatt gtggcgctgc ttgcgagaac 120  
acgtactctg ttgacggaaa ggaggttaca aaattctctt gtaaccaaca gtcgggaaga 180  
tgtgccagga gtgtttatag tgcgtcttct ctagaaagag cagccaatga tcttgccac 240  
attataggta tcatcaagaa aaatccaaaa ttggaggaag aactccctga atcatttttg 300  
tggtttatca atcacaatgg aggagatttg tttgtgaata agcgagccgc gtactacgac 360  
acgatgcac taagcatagg gaaactggat aatgtggaca ctcttgccca gggtttagat 420  
aaacggatgg cctcatcatt gagagagcac ctactgagga agttggactc tatactttta 480  
caaattgata aagttaaata tgaaaaggca aagaaatgga tattggatat aacacaggag 540  
gctggcaccg aagaggacaa taaagaagaa gaagatgcga aaaaagagga tcaatctctt 600  
agcgtttagt aaattgtgga tgttttaacg ggcacacatg accctatgcc cctgagggct 660  
agagggttta tccagaaaaa aatataatcct ttgtcaagaa acgaattgag agaattagcc 720  
cttaaggaac ttttccctga agaaactaca tctcctcagg ttttgagtag gcaacatgat 780  
gtatctacgc gtgaagattt atgcaatgaa agtatgaatg cagggagggc agaattccatt 840  
tttagcgacc ctgattcttg agagtacgtg gctacttgtg catgtcttta ctcggaatat 900  
ttaacagggc ctgctgtgta gcacaaaaca tacaggtagg ttatagacta cgacaaatgg 960  
aaaaggactg gaagacctga atttctaact gatcctgtac ttcattttta aaaggcagaa 1020  
gctgtgtgta aatcgacaaa tccaaacttg agggcaattt atagtccaga taataaagg 1080  
ttcttgtgtg cgctgttagc tgaacttgta aagacggcat taacttttag gggttcacac 1140  
gaaccgtctc tcattgtcga gagagatata aatcaagctg aaaatctacc atccaattca 1200  
tttgggtgaa actggcccta tgtgaatctc ctaaactgca ttcaagacca gtacacgtaa 1260  
tttgataa 1268

<210> 51  
<211> 413  
<212> PRT  
<213> SHRIMP

<400> 51  
Met Ser Ser Thr Asp Leu Ser Lys Asn Ala Phe His Asp Trp Val Val  
1 5 10 15  
Ser Lys Thr Asp Cys Glu Val Phe Asp Val His Cys Glu Thr Asp Arg  
20 25 30  
Asp Cys Gly Ala Ala Cys Glu Asn Thr Tyr Ser Val Asp Gly Lys Glu  
35 40 45  
Val Thr Lys Phe Ser Cys Asn Gln Gln Ser Gly Arg Cys Ala Arg Ser  
50 55 60  
Val Tyr Ser Ala Ser Ser Leu Glu Arg Ala Ala Asn Asp Leu Gly His  
65 70 75 80  
Ile Ile Gly Ile Ile Lys Lys Asn Pro Lys Leu Glu Glu Glu Leu Pro  
85 90 95  
Glu Ser Phe Leu Trp Phe Ile Asn His Asn Gly Gly Asp Leu Phe Val  
100 105 110  
Asn Lys Arg Ala Ala Tyr Tyr Asp Thr Met His Leu Ser Ile Gly Lys  
115 120 125  
Leu Asp Asn Val Asp Thr Leu Ala Gln Gly Leu Asp Lys Arg Met Ala  
130 135 140  
Ser Ser Leu Arg Glu His Leu Leu Arg Lys Leu Asp Ser Ile Leu Leu  
145 150 155 160  
Gln Ile Asp Lys Val Lys Tyr Glu Lys Ala Lys Lys Trp Ile Leu Asp  
165 170 175

Ile Thr Gln Glu Ala Gly Thr Glu Glu Asp Asn Lys Glu Glu Glu Asp  
 180 185 190  
 Ala Lys Lys Glu Asp Gln Ser Leu Ser Val Ser Glu Ile Val Asp Val  
 195 200 205  
 Leu Thr Gly Thr His Asp Pro Met Pro Leu Arg Arg Phe Ile Gln Lys  
 210 215 220  
 Lys Ile Tyr Pro Leu Ser Arg Asn Glu Leu Arg Glu Leu Ala Leu Lys  
 225 230 235 240  
 Glu Leu Phe Pro Glu Glu Thr Thr Ser Pro Gln Val Leu Ser Arg Gln  
 245 250 255  
 His Asp Val Ser Thr Arg Glu Asp Leu Cys Asn Glu Ser Met Asn Ala  
 260 265 270  
 Gly Arg Ala Glu Ser Ile Phe Ser Asp Pro Asp Ser Gly Glu Tyr Val  
 275 280 285  
 Ala Thr Cys Ala Cys Lys Glu Tyr Leu Thr Gly Pro Ala Cys Lys His  
 290 295 300  
 Lys Tyr Tyr Val Ile Asp Tyr Asp Lys Trp Lys Arg Thr Gly Arg Pro  
 305 310 315 320  
 Glu Phe Leu Thr Asp Pro Val Leu His Phe Lys Lys Ala Glu Ala Val  
 325 330 335  
 Cys Lys Ser Thr Asn Pro Asn Leu Arg Ala Ile Tyr Ser Pro Asp Asn  
 340 345 350  
 Lys Gly Phe Leu Cys Ala Pro Val Ala Glu Leu Val Lys Thr Ala Leu  
 355 360 365  
 Thr Phe Arg Gly Ser His Glu Pro Ser Leu Ile Val Glu Arg Asp Ile  
 370 375 380  
 Asn Gln Ala Glu Asn Leu Pro Ser Asn Ser Phe Gly Val Asn Trp Pro  
 385 390 395 400  
 Tyr Val Asn Leu Leu Asn Arg Ile Gln Asp Gln Tyr Thr  
 405 410

<210> 52  
 <211> 1401  
 <212> DNA  
 <213> SHRIMP

<400> 52  
 atgtctgcat ctttaatat ggacgaatac ctcaagaaga ctgcttcagc cggtctggat 60  
 gtagctgact catttgagaa aatcaaggga gaaatccaat cacctgagga ggctgcggct 120  
 ctttctgttg ctctctatgg agcacctcca aaaccttcag cttcggtgtt ggctctctatc 180  
 atcactggag aaagaacatc tttaaagcag aaatatctat cggataatgt cctattgaaa 240  
 atgtctgttg ctgcggttg acaagaaaat aatcgcaaga gagccgacca ggagctgat 300  
 gaaattagaa ccatcatgga agatattaca gggagtttgt ccggtgcgta caggcaatat 360  
 agcccgcctc aggaagaaaa taaggtgcat ataggcatca tgaataacaa aacgcctagc 420  
 attgtttgtg gatattatac aatggacaca tctattttct ccgaacctct ttctctaaca 480  
 gattttcaaa accccactgt cattgccaat gtgactaagc ggatggagag cattttttca 540  
 aaggtcgact ctgctagggt tacaagattc gacgcttttg ttaatggtgt tgcgaataat 600  
 atggatataa agtcatcaat agattgggca aatatggtag aaaatgtgat caaattacca 660  
 gattctacac ctaacccttg ttcagttgac actattgtgt ccagagacgc aagtgtagtt 720  
 aaaacagcag ttaatgatat atacgcttct gttggaaaat cttattgtcg tcttgcaaca 780  
 cagctaacct ttatgagcga gattgaaaaa ctgcgaaagg ctgcagttgt atgttttgag 840  
 gcactcatgt ccgatactag ggagaggcca ttcgtagagt tcctatttta cgtagcttt 900  
 aaggaagatg catcaaatc caattcaaaa ttgtttgttc agaataagct atcttccatg 960  
 tctggaaacc ccagacagcc cataaaattg gtacgccgtt ctgctgagga aacactattc 1020  
 gggctctgtt tcatgtttta ggtaatgcct ccagaattca tgaactgtat atttaacttc 1080  
 cctaccattc cccattcaac acaataccat ggtctatatg gtacatgttt aacccctcta 1140  
 cttagaaaat acggttcttc attcgaaaag tctctgggctc attttgagga aattttaagc 1200  
 gaaagagcca atgcagtgaa aaaatttggt gtaaacgata cgaggataga ttgtctagat 1260  
 gcagtagcaa atctcaccgg acctgtgtat gttctcattt tagatcttgt acgtactcta 1320  
 agtgcgcaga gatcgtgttc aactaaattt ctccgtgaaa ttaaggaaaa ctatcttttg 1380

WO 01/38351

132

PCT/US00/28888

tggaataggt ttgtgtcata a

1401

<210> 53

<211> 459

<212> PRT

<213> SHRIMP

<220>

<221> VARIANT

<222> (1)...(459)

<223> Xaa = Any Amino Acid

<400> 53

Met	Ser	Ala	Ser	Leu	Ile	Leu	Asp	Glu	Tyr	Leu	Lys	Lys	Thr	Ala	Ser
1				5					10					15	
Ala	Val	Leu	Asp	Val	Ala	Asp	Ser	Phe	Glu	Lys	Ile	Lys	Gly	Glu	Ile
			20					25					30		
Gln	Ser	Pro	Glu	Glu	Ala	Ala	Ala	Leu	Ser	Val	Tyr	Gly	Ala	Pro	Pro
		35					40					45			
Lys	Pro	Ser	Ala	Ser	Ala	Val	Ala	Ser	Ile	Ile	Thr	Gly	Glu	Arg	Thr
	50					55					60				
Ser	Leu	Asn	Asp	Lys	Tyr	Leu	Ser	Asp	Asn	Val	Leu	Leu	Lys	Met	Ser
65					70					75				80	
Val	Ala	Arg	Val	Gly	Gln	Glu	Asn	Asn	Arg	Lys	Arg	Ala	Asp	Gln	Ala
				85					90					95	
Ala	Asp	Glu	Ile	Arg	Thr	Ile	Met	Glu	Asp	Ile	Thr	Gly	Ser	Leu	Ser
			100					105					110		
Gly	Ala	Tyr	Arg	Gln	Tyr	Ser	Pro	Leu	Glu	Glu	Glu	Asn	Lys	Val	His
		115					120					125			
Ile	Gly	Ile	Met	Asn	Asn	Lys	Thr	Pro	Ser	Ile	Val	Cys	Gly	Tyr	Tyr
	130					135					140				
Thr	Met	Asp	Thr	Ser	Ile	Ser	Ser	Glu	Pro	Leu	Ser	Leu	Thr	Asp	Phe
145					150					155					160
Gln	Asn	Pro	Thr	Val	Ile	Ala	Asn	Val	Thr	Lys	Arg	Met	Glu	Ser	Ile
				165					170					175	
Phe	Ser	Lys	Val	Asp	Ser	Ala	Arg	Ser	Thr	Arg	Phe	Asp	Ala	Phe	Val
			180					185					190		
Asn	Gly	Val	Ala	Asn	Asn	Met	Asp	Ile	Lys	Ser	Ser	Ile	Asp	Trp	Ala
		195				200						205			
Asn	Met	Val	Glu	Asn	Val	Ile	Lys	Leu	Pro	Asp	Ser	Thr	Pro	Asn	Pro
	210					215					220				
Cys	Ser	Val	Asp	Thr	Ile	Val	Ser	Arg	Asp	Ala	Ser	Val	Val	Lys	Thr
225					230					235					240
Ala	Val	Asn	Asp	Ile	Tyr	Ala	Ser	Val	Gly	Lys	Ser	Tyr	Cys	Arg	Pro
				245					250					255	
Ala	Thr	Gln	Leu	Thr	Phe	Met	Ser	Glu	Ile	Glu	Lys	Leu	Arg	Lys	Ala
			260					265					270		
Ala	Val	Val	Cys	Phe	Glu	Ala	Leu	Met	Ser	Asp	Thr	Arg	Glu	Arg	Ala
		275					280					285			
Phe	Val	Glu	Phe	Leu	Phe	Tyr	Val	Ser	Phe	Lys	Glu	Asp	Asn	Thr	Asn
	290					295					300				
Ser	Lys	Leu	Phe	Val	Gln	Asn	Lys	Leu	Ser	Ser	Met	Ser	Gly	Asn	Pro
					310					315					320
Arg	Gln	Pro	Ile	Lys	Leu	Val	Arg	Arg	Ser	Ala	Glu	Glu	Thr	Leu	Phe
				325					330					335	
Gly	Leu	Cys	Phe	Met	Phe	Lys	Val	Met	Pro	Pro	Glu	Phe	Met	Asn	Cys
			340					345					350		
Ile	Phe	Asn	Phe	Pro	Thr	Ile	Pro	His	Ser	Thr	Gln	Tyr	His	Gly	Gly
		355					360					365			
Thr	Cys	Leu	Thr	Pro	Leu	Leu	Arg	Lys	Tyr	Gly	Ser	Ser	Phe	Glu	Lys
	370					375					380				



Met 1	Ala	Gln	Thr	Ser 5	Lys	Met	Gly	Thr	Asn 10	Lys	Arg	Cys	Phe	Glu 15	Glu
Glu	Val	Glu	Glu	Glu	Arg	Gln	Gln	Pro 25	Phe	Thr	Lys	Lys	Ser 30	Lys	Ser
Glu	Pro	Pro 35	Ser	Phe	Glu	Asp	Lys 40	Ser	Ser	Ser	Thr	Ser	Ser	Lys	Lys
Lys	Ser	Lys	Ser	Asn	Lys	His 55	Thr	Lys	Thr	Lys	Glu	Glu	Gln	Leu	Leu
Glu 65	Phe	Val	Lys	Asp	Leu 70	Glu	Arg	Ser	Asp	Pro 75	Thr	Val	Pro	Asp	Glu
Lys	Val	Lys	Gln	Glu	Val	Glu	Glu	Lys	Ser 90	Pro	Glu	Ala	Ile	Ala	Glu
Ile	Phe	Ser	Met 100	Phe	Gly	Ile	Ala	Gln 105	Asp	Ser	Lys	Phe	Lys	Ser	Leu
Leu	Pro	Ile 115	Glu	Arg	Ile	Lys	Ser 120	Ile	Thr	Thr	Lys	Ile	Val	Ile	Asp
Ala	Ile 130	Asn	Gln	Pro	Val	Arg 135	Lys	Met	Leu	Val	Asp 140	His	Leu	Tyr	His
Phe 145	Lys	Glu	Met	Gln	Asn 150	Val	Val	Glu	Lys	Tyr 155	Lys	Asp	Asp	Ser	Asp
Glu	Lys	Leu	Ser	Val 165	Ile	Leu	Lys	Ser	Lys 170	Lys	Ser	Pro	Lys	Glu	Phe

WO 01/38351

134

PCT/US00/28888

Asp	Leu	Ser	Phe	Ser	Asp	Tyr	Val	Asp	Arg	Leu	Asn	Arg	Ile	Leu	Val
			180					185					190		
Gly	Val	Ile	Lys	Arg	Val	Ala	Gly	Ala	Ile	Glu	Ser	Lys	Glu	Leu	Leu
		195					200					205			
Gln	Ser	Asn	Ser	Met	Ile	Met	Asn	Ser	Val	Leu	Gly	Thr	Val	Val	Ser
		210				215					220				
Asn	Ile	Pro	Tyr	Asn	Met	Lys	Ile	Asn	Ile	Cys	Val	Phe	Leu	Thr	Asn
225					230					235					240
Phe	Ile	Cys	Thr	Phe	Ala	Asn	Asp	Asp	Leu	Tyr	Thr	Phe	Phe	Arg	Asp
				245					250					255	
Asp	Glu	Lys	Phe	Val	Met	Ser	Gln	Val	Thr	Arg	Tyr	Ile	Ser	Lys	Asp
			260					265					270		

<210> 56  
 <211> 1398  
 <212> DNA  
 <213> SHRIMP

<400> 56

atggctgtag	gggattatct	ttctatgtcc	tctgtagggg	aagcaacgct	cgtaggggtttt	60
atgatttttaa	attttataaaa	ctttgtcaca	atattatctt	taataatata	tgccgtgaca	120
gacgtttata	gaagatgcaa	gcggccgtct	accaatggat	attctgggtt	tacaaccaat	180
gtggtttctt	cgactttaca	agaagcta	ctcgttacca	ctgaaaaaga	taaacctgtt	240
caatttggtta	gagggttagt	ccccagaaaa	atgatggaaa	aatatagatc	ggacttgtct	300
cctaagaacg	tgggggaata	tattttacct	tcagaaaaag	aaacagacaa	attgaaaagt	360
gattataaaa	agggtaagaa	ggttggtctt	ttaactgccc	tgagtaatgg	tcatgacagt	420
aacaagagga	ttatagggcc	aagggatctg	attagtagag	atgatgtgaa	ggacaaaagt	480
tatgtcttta	agagattgag	caaagatccg	ctcgtctact	actcttctgc	aacctctaaa	540
tacgttagaa	aattttcccc	tttcagagca	aaaaaattca	tgacatcaac	acagttgggg	600
agtaagctcg	tgtatcctca	ccctatacgg	tatggtactg	cttttgtact	acccacggga	660
tacgtgatca	acaaagcata	cggaaatgat	aatgaggatc	tacacacttg	gaaccacccc	720
tcttcttctg	tgctcgtccc	agactcta	aatgatagat	taacagtaga	atgtgctaaa	780
acggacccaa	cacatagaat	cggcatctac	ggctttggag	ggtctgatga	taatagacgt	840
gcaaaagaag	aaggttatgt	agaaatgttg	ttatgtaatt	gtgacaacca	caaggacttg	900
cttaaggctc	ctctaattac	agagtattct	acaaatccaa	ctgaaattca	agtagatgtt	960
gctgcaaaac	gtgtttttatt	ccctgcccct	ggttccgagc	ctgtaaaatc	ttcccaagtg	1020
acatctgctg	ctcatcaact	agacggagct	actggcgagc	acgatatttc	ccatgagccc	1080
gtgaagctat	cagatacggg	tgactatgca	gttggtatcac	ccattgtatt	caagccagtt	1140
tatggtacat	cttttagtaaa	tcttccagaa	acaggatctc	ctctggcatt	gaactgcccc	1200
tgacccgaca	aggctgatgg	aatatatcaa	gtcaatcaaa	agggagggat	attatataga	1260
gatatggtgg	ggtatcttaa	cgccaaccct	gtggaagctg	catcactttc	ttcctcggac	1320
tcttcttcgt	ggttgacaac	tggtaacaaa	atatcttctg	ttacatgtga	aggagaaaaa	1380
ataaagaaaa	ttgtgtaa					1398

<210> 57  
 <211> 463  
 <212> PRT  
 <213> SHRIMP

<400> 57

Met	Ala	Val	Gly	Asp	Tyr	Leu	Ser	Met	Ser	Ser	Val	Gly	Glu	Ala	Thr
1				5				10						15	
Leu	Val	Gly	Phe	Met	Ile	Leu	Asn	Phe	Ile	Asn	Phe	Val	Thr	Ile	Leu
			20					25					30		
Ser	Leu	Ile	Ile	Tyr	Ala	Val	Thr	Asp	Val	Tyr	Arg	Arg	Cys	Lys	Arg
		35					40					45			
Pro	Ser	Thr	Asn	Gly	Tyr	Ser	Gly	Cys	Thr	Thr	Asn	Val	Val	Ser	Ser
	50					55					60				
Thr	Leu	Gln	Glu	Ala	Asn	Leu	Val	Thr	Thr	Glu	Lys	Asp	Lys	Pro	Val
65					70					75				80	

```
<400> 58
atggagtaca  tggaagaag  agacatcgct  gaaaggcgct  cagaaggtgt  cgactatatt  60
ctggacgaaa  actctgcttg  tgtagttaat  gtgaagagta  tccgtaacag  gctcggtgcc  120
atggacgccg  aggaggcaca  gtagcgcacg  gacatttccg  cccaacttgc  caccatatt  180
atccgtctgg  cccactgctc  cgaatccaac  aagattaagg  ataccattgc  cagtattgcg  240
ggtcttttca  tcaacaacat  ctttgacaac  aattcaacaa  agaacaaact  taaaacqtat  300
```

WO 01/38351

136

PCT/US00/28888

```

aatcaattca aggcagagtc acaaaacaag tctagcggtc tcaatatctt tggctctcta 360
gatcctctga gtatgctttc tagcttcatg ggttctgata cagcaaagag tggaggggaa 420
aatttggaca aatctttggg tgtgctcttt gaggtgcttc aaaattacaa cccttgcaag 480
attgacgata ttgtccttct ggaaatgtgc ccatccaagt gcgcgcgctg caccgggtctc 540
aaggaagcca tccgccagga acaacccatg gaagcaatgt tgttggtttt caaatgtatc 600
aaccataata ggttcaattt tggaaagcag ataaagtcag catacgctc tgaaacatgc 660
atgagatact ctcaggacga acgcgcagtc gttgtgcctc tgaggagtat cctcctcggc 720
tgccctcgaca gggatgatcc agctcatact ctttctctct tcggggatac tatcgagtat 780
gctgattcag ataacgcttg ggtttcttcc ctggttgagc ccgtctctag aatgcctatg 840
gtagacagag ccgttattgc tcacttttac gtgtacacaa tgctcagccg acataggcga 900
gtatctggag acagcttcaa gcagtttgtc tataccgtat ttgttcgtat gatttactct 960
gcgattgaaa ttttggtctg tgatactgaa aattcgtctg tagaatgtga tggaaagcac 1020
tttttgagct atgttaatgc catggttaac gtatccgtgc tgggttctac gtttaacgta 1080
ctaaaagcct accgttcatg ggtggtggat caagcatccg tcgcaccggt tctagacatt 1140
atttccggag gatggaagaa gaactacccc tcacctgacc acatcaagag ggtggcgta 1200
gacatctctc aagtcaccaa tcactcttga tcaccttcta gaatgggtta aggtaacaac 1260
aaggctagca acgttacatc tggcctggat agtatcaggt ctgttcgtca agcagaaaaa 1320
tatatcccgt ttggaatact tgaaaataaa gcagggtatg gtgtaataaa cattgccaa 1380
cacaacatca gccgtccagc aagagaacaa tccaacggaa ggaactttaa ttgcaacgct 1440
ttacacattc taccttcaat taagggtctg gaggcacttg gggcacaaaa ggggagcgca 1500
gatcaaacctg tcaatgtttt tgataatttt gtcgcatctc atatggatat tgccatgaaa 1560
aagcaggggt cggggaagat tcttggactg ctactagca tgattgacag gcaaggctctg 1620
actacttcat tccctagtag tgaagcggaa tacaagaaga gaatccatga tttcacaaga 1680
tacgtcatct tctcttcaac acccatcaac gacgaactag tcaattctcg ctgtattctt 1740
ccccattcta atgttctgaa ctccccctatc agcttgagaa atattgacct agaatcagtc 1800
cccgacactc gattccactt tctgctcatg atgtggcagc ggccaaatat cgtgaacct 1860
aatctttctg ctctaactac aagtcagcta gaattgttgc ttagcaagaa ccaaaaaatg 1920
gacaaactca ccaccagagc gttcttcaat atcgacagga tcaatttcca gatggcagac 1980
gctatcatta agaacgtttc tggaaagcggc ttcctagatg ggagtaaaac tgcctcttct 2040
tcttctctcag cgcctaactt tttccaaatc ttcagtggtg ctgaatgcac tgcaaagcag 2100
ctccaaagta ttcgcaaatt cattggagaa tctatgcagc atgtacaaaa ggaatggagt 2160
agtgcagtaa acaatgggaa cagaggagta gaaaattatg acggactcaa tgctcagttc 2220
tctgaagaac tgttcgagct gctctacaaa ttgatcatcg aggaggatat gcggccatcc 2280
agcctgatcg cctcatctga attcttgagc aactacgtca acgcatgga tgaacttctt 2340
atcagagcta atgtctctta g 2361

```

<210> 59  
 <211> 778  
 <212> PRT  
 <213> SHRIMP

```

<400> 59
Met Glu Tyr Met Glu Glu Gly Asp Ile Ala Glu Arg Arg Ser Glu Gly
 1          5          10          15
Val Asp Tyr Ile Leu Asp Glu Asn Ser Ala Cys Val Val Asn Val Lys
 20          25          30
Ser Ile Arg Asn Arg Leu Gly Ala Met Asp Ala Glu Glu Ala Gln Tyr
 35          40          45
Ala Gln Asp Ile Ser Ala Gln Leu Val Thr His Ile Ile Arg Leu Ala
 50          55          60
His Cys Ser Glu Ser Asn Lys Ile Lys Asp Thr Ile Ala Ser Ile Ala
 65          70          75          80
Gly Leu Phe Ile Asn Asn Ile Phe Asp Asn Asn Ser Thr Lys Asn Lys
 85          90          95
Leu Lys Thr Tyr Asn Gln Phe Lys Ala Glu Ser Gln Asn Lys Ser Ser
100          105          110
Val Leu Asn Ile Phe Gly Ser Leu Asp Pro Leu Ser Met Leu Ser Ser
115          120          125
Phe Met Gly Ser Asp Pro Ala Lys Ser Gly Gly Glu Asn Leu Asp Lys
130          135          140
Ser Leu Gly Val Leu Phe Glu Val Leu Gln Asn Tyr Asn Pro Cys Lys

```

145 150 155 160  
 Ile Asp Asp Ile Val Leu Leu Glu Met Cys Pro Ser Lys Cys Ala Ala  
 165 170 175  
 Cys Thr Gly Leu Lys Glu Ala Ile Arg Gln Glu Gln Pro Met Glu Ala  
 180 185 190  
 Met Leu Leu Phe Phe Lys Cys Ile Asn His Asn Arg Phe Asn Phe Gly  
 195 200 205  
 Ser Asp Ile Lys Ser Ala Tyr Ala Ser Glu Thr Cys Met Arg Tyr Ser  
 210 215 220  
 Gln Asp Glu Arg Ala Val Val Val Pro Leu Arg Ser Ile Leu Leu Gly  
 225 230 235 240  
 Cys Leu Asp Arg Asp Asp Pro Ala His Thr Leu Ser Ser Phe Gly Asp  
 245 250 255  
 Thr Ile Glu Tyr Ala Asp Ser Asp Asn Ala Trp Val Ser Ser Leu Phe  
 260 265 270  
 Ala Ala Val Ser Arg Met Pro Met Val Asp Arg Ala Val Ile Ala His  
 275 280 285  
 Phe Tyr Val Tyr Thr Met Leu Ser Arg His Arg Arg Val Ser Gly Asp  
 290 295 300  
 Ser Phe Lys Gln Phe Val Tyr Thr Val Phe Val Arg Met Ile Tyr Ser  
 305 310 315 320  
 Ala Ile Glu Ile Leu Phe Cys Asp Thr Glu Asn Ser Ser Val Glu Cys  
 325 330 335  
 Asp Gly Lys His Phe Leu Ser Tyr Val Asn Ala Met Val Asn Val Ser  
 340 345 350  
 Val Leu Gly Ser Thr Phe Asn Val Leu Lys Ala Tyr Arg Ser Trp Val  
 355 360 365  
 Val Asp Gln Ala Ser Val Ala Pro Val Leu Asp Ile Ile Ser Gly Gly  
 370 375 380  
 Trp Lys Lys Asn Tyr Pro Ser Pro Asp His Ile Lys Arg Val Ala Tyr  
 385 390 395 400  
 Asp Ile Ser Gln Val Ile Asn His Leu Asp Ser Arg Met Val Lys Gly  
 405 410 415  
 Asn Asn Lys Asn Val Thr Ser Gly Leu Asp Ser Ile Arg Ser Val Arg  
 420 425 430  
 Gln Ala Glu Lys Tyr Ile Pro Phe Gly Ile Asn Lys Ala Gly Tyr Gly  
 435 440 445  
 Val Ile Asn Ile Ala Lys His Asn Ile Ser Arg Pro Ala Arg Glu Gln  
 450 455 460  
 Ser Asn Gly Arg Asn Phe Asn Cys Asn Ala Leu His Ile Leu Pro Ser  
 465 470 475 480  
 Ile Lys Gly Cys Glu Ala Leu Gly Ala Gln Lys Gly Ser Ala Asp Gln  
 485 490 495  
 Thr Val Asn Val Phe Asp Asn Phe Val Ala Ser His Met Asp Ile Ala  
 500 505 510  
 Met Lys Lys Gln Gly Ser Gly Lys Ile Leu Gly Leu Thr Ser Met  
 515 520 525  
 Ile Asp Arg Gln Gly Leu Thr Thr Ser Phe Pro Ser Ser Glu Ala Glu  
 530 535 540  
 Tyr Lys Lys Arg Ile His Asp Phe Thr Arg Tyr Val Ile Phe Ser Ser  
 545 550 555 560  
 Thr Pro Ile Asn Asp Glu Leu Val Asn Ser Arg Cys Ile Leu Pro His  
 565 570 575  
 Ser Asn Val Leu Asn Ser Pro Ile Ser Leu Arg Asn Ile Asp Pro Glu  
 580 585 590  
 Ser Val Pro Asp Thr Arg Phe His Phe Leu Leu Met Met Trp Gln Arg  
 595 600 605  
 Pro Asn Ile Asp Glu Pro Asn Leu Ser Ala Leu Thr Thr Ser Gln Leu  
 610 615 620  
 Glu Leu Leu Leu Ser Lys Asn Gln Lys Trp Asp Lys Leu Thr Thr Arg  
 625 630 635 640

Ala Phe Phe Asn Ile Asp Arg Ile Asn Phe Gln Met Ala Asp Ala Ile  
 645 650 655  
 Ile Lys Asn Val Ser Gly Ser Gly Phe Leu Asp Gly Ser Lys Thr Ala  
 660 665 670  
 Ser Ser Ser Ser Ser Ala Pro Asn Phe Phe Gln Ile Phe Ser Gly Ala  
 675 680 685  
 Glu Cys Thr Ala Lys Gln Leu Gln Ser Ile Arg Lys Phe Ile Gly Glu  
 690 695 700  
 Ser Met Gln His Val Gln Lys Glu Trp Ser Ser Ala Val Asn Asn Gly  
 705 710 715 720  
 Asn Arg Gly Val Glu Asn Tyr Asp Gln Ala Gln Phe Ser Glu Glu Leu  
 725 730 735  
 Phe Glu Leu Leu Tyr Lys Leu Ile Ile Glu Glu Asp Met Arg Pro Ser  
 740 745 750  
 Ser Leu Ile Ala Ser Ser Glu Phe Leu Ser Asn Tyr Val Asn Ala Met  
 755 760 765  
 Asp Glu Leu Leu Ile Arg Ala Asn Ala Ser  
 770 775

<210> 60  
 <211> 18234  
 <212> DNA  
 <213> SHRIMP

<400> 60  
 atggaccagt acccagaagt gagggatact cctcagacag aacaggaaca ggcggcggcg 60  
 caacaacaac aagcagcaac aacaacagct gccgctgctg ctgccgctgc tcctacgcag 120  
 tacagtaaca ctgtttctgc agaaacttta tccgccattt ctgaagatgg aaaattggag 180  
 aggtcaatcg cagcttcgtg ctggatcaat aaccttaacc ctgatgaaaa aatggctcaa 240  
 cgtgtccaat tccatccact tagttccacg accacatacg attcagaaaa tgtgaacctt 300  
 ggtagtctcg ttgtgttttt gaagcctaga gccctcccca ccggggggcac gtgttttagc 360  
 cccaactaca ttgcagtgcc tactcttcgt gctgcgtcag aaattatcga ttctattgca 420  
 tcaactagtc tataccaatg ttcaatgttc aattcgtgga atcttattcc catttttatg 480  
 agtaatagca aacatagtc attcggatgat cgtgtcatca agagatcaat gatccgtaat 540  
 tgtttttcca aacaaaaaaaa tgtagaaaaa cttttaaagg aattgcgcag gagaaagggtt 600  
 aacgccgcca aggcattttc tcatgcagtt caacagaaat cagcagtaaa tactgccctt 660  
 gcagcatgga acgcgggaag tgctgcaaac ctcgaaaaat tagtagattt ttgcaaatg 720  
 aagtacagtc ccgatagaaa atacaaagca ggaggtcttt ttagtgctgc agcaacagct 780  
 cagtacacaat caggtacttc ttcttcctca gttgaacata cttctaataa ttcttactc 840  
 gatattcctaa aacgtcaciaa gggaacgtcc cttgatttag attcagcaac aaacacattt 900  
 gatactgctc ttctcagggt tttaccgag aggttagagc agccgtagat 960  
 gcagccgcag attcagatca cctatccgcg tcggacccta ttttctctat ttttcgcac 1020  
 aatagcaggc gtgagggtat tctagattca gtgccaaata tcggtatgtt ggctccacga 1080  
 tcaaaatatt ctgtcgcaga gtacttgatg gcagacagag acgagtctgc agatattcgt 1140  
 gctaagatag ggacaaagat tgcaacagat tttgaagcat tgagaggaga taataataag 1200  
 agaagagcag atacttcagt cgatgatctc aaagaatcac tagcggattc tattgaaaaa 1260  
 acttccatta aaaataccgg cgatattaat tcggtcacaa atattcccac agatacagag 1320  
 gagtatgaat tttctctcca catcacacag cttttcgcgc aggcattttt ggaaactatg 1380  
 ggaagtttat tgagctgtgc tttcgggtga cagttcccct tctcagatga aggttttgcg 1440  
 gctatcgaaa ggattatacg taaaacagat cctgatactg gaaaggtttc agaaatggac 1500  
 ccttctctc tttcagatca atatctcttg ttggtaggaa atttccagggt ttctcccttc 1560  
 cacgtttcag atoctaagga tatcgttttc ggtagacaag ttacgcctaa tacgcctatt 1620  
 ttgagtatca ttactagatc taaaaatgat aagaatgaaa cgtctactat tattaatttc 1680  
 agggatagcg ttcttggttaa cgataccgtg ttgagagacg ccacgcaaaa cgtctcaact 1740  
 tccaccctt ctcaaagaag agtccccacc gccgcagggtg aaccaaaaaa gccaatgctc 1800  
 tctggttgcc tacctatcat caggggaccc caggtagtga ccagggaaaq tgatgatattg 1860  
 attagtgtc ttgtgggtga ttgttatatc tctttggggg tctattatgc catgggttca 1920  
 tcggctgcag ctattgcccg tggccaccag agagctctcg cttcagccga gtcaattaat 1980  
 tcacccatga tgaaaaaatt ctcaaagaag ggaggaaagt atacagagga ggaaaagcga 2040  
 atcaagaagg ctatgcgtag gaatgcagat cgttctgcta ggattttggc tttgttgggg 2100

caaaactgatg	cccagtatgg	gtatgtcgaa	cataattcta	ccctagattc	gttctggtct	2160
tcgaatgccg	ctattcgtgc	aaaggcaaag	gaggatgctc	ttagccgtgc	agaaatcttg	2220
gcagttagga	aacaattgga	cggaaaatgt	tcctcttcaa	gggatgaata	ttccatggta	2280
gagagatate	ttagagactc	atttttcagg	tcagttaata	ggtcaggagg	aggatatgaa	2340
atgttttgatc	aaggttttga	tatgggaaga	tttccgact	ttttgagtga	caactctgct	2400
gccaggaacg	cttggcaaca	gtacgcagaa	gtaatgagag	gactttctaa	gcattgagaaa	2460
cgcgtattca	acattgaagg	tcttttcagt	gctctaaatt	ctttcaagtt	cccccttggt	2520
ccagaacagg	ggcgtaaaaa	gactgtaggt	ggaaggcata	ggcttaacaa	tttgaaggcg	2580
gccaaataaga	tcattaacgg	catcacagag	atgactctcc	agtcagccat	cgatggtact	2640
ggaatctctg	atattattgg	ttcagtttagc	gatgggtggg	gaaatactac	agctcagccg	2700
tcctcgcgtca	aagctcttaa	aacattgtcc	aatttcagcg	gaaacggaaa	tgttgtgtca	2760
attccagtta	gtcagagcagt	aaagtgtgcc	gcgggaagtc	gaggtgggga	aactctcaag	2820
tgtgtggaca	ttccttcagt	cattattgca	aacctaattct	ctgataagag	aattctagat	2880
caactttgcg	gaggagggat	gaacctcgct	cacgaaatca	caaactttat	cgagacgatt	2940
gcaggtaaaag	aacatacagg	aaaggaatcc	gttttcttgt	ctcctagatt	gtctgtcatt	3000
ctttttgcgtt	atatttggtt	caacgcagcg	gttgtttccc	ttacggatag	caacattaaa	3060
atgccccctca	acacaatgtc	cgagggcact	ggcgatgaca	tttatagggga	ctattttggcc	3120
atcagaggca	tggtaataaa	ttacaacagt	agtctctcct	ctattttcagt	caaggctatt	3180
tcggataggt	acaattgctg	tagtggaat	acttcaacca	gtaataagaa	tgtgaccatt	3240
aaagactcaag	gtgaattggt	aactgtcctt	caacagactg	ctaattgcctt	gtccgccttt	3300
accaacaagg	gagcggtggg	tgcaaccccc	gatgtgcca	atatggccaa	cgttatttcc	3360
ccaattgcta	atgcggatgt	agttaagaac	accaacgtgg	ttgtttcagg	gctagatagg	3420
atcactgaga	ccatcaactt	cttttcattt	ttgtctcaga	tcaaaacaat	gaacgagAAC	3480
attgaagagt	atcttaggag	atataggcta	ggagaaggac	tagataagaa	agaattggat	3540
aatttttgtgt	atccaaatat	tgcagctatt	gttaagcgag	aattgggggt	aagtgggtcc	3600
gcattgtcca	gtaatctoga	tactgactgt	ccaattacta	togacctgaa	caactgaacag	3660
cctttgatcg	taaaggctag	caagggttat	gcctctaacc	gctacgctaa	attattcaac	3720
aaaacaacaa	gaacagcagc	agaacaagct	cagatggagc	agtataatgc	acaaatggct	3780
gccaaatacta	ttcctcaatt	agtaaacagg	ttgaccatcc	ctggatccat	cacggcagac	3840
actgccatca	atgtcggttaa	agctttcaca	gaaaatggag	aatttagtaa	cgcagaaaca	3900
cacttggggg	ttatgggtaa	cgcgattaat	gaaatgcaac	ctcttttcac	ggacggattc	3960
aacgtttgcga	acaagcggtt	aacagttaac	gtgggttcag	ttagtaagct	gattcagaat	4020
gggttaaccg	tatctctcat	tcttgctcac	tcaaaggcta	gcccctatgt	ctttaagcct	4080
ctcgtgcaag	atttcgctaa	gcttttactg	gcagtcactg	cagagacttc	tctggttgct	4140
tctaggtccc	agaagagttt	cttccccatt	cctccttcag	tattttcttc	aggtgggtctt	4200
ttcaaaattg	atagggaaat	gttcgataat	atgaaagacag	attatgtagt	ggaagtaatt	4260
agacagctat	ctaagaatgc	taccgccgcc	atcgaaaagg	gcaatgattc	cgattcagct	4320
gctaggattg	ccaagtcagg	tgaattttat	aacaaggatg	ttgcatcaac	caactgcagct	4380
cccggaaactt	cttctccgcg	tttaaccttg	ttcgccaata	atctccagaa	ccctgcaaaag	4440
gtatgggtcta	tgggagctct	cccccatctt	gatattggccg	tggtaccaaaa	acttcatggc	4500
atttctcacg	atcaaatggt	ccgcctatct	acataattatc	agggatttca	taagatggaa	4560
cttaacagcg	attgcaaaacc	agaagaatgc	gataattctc	ttcctggaaa	tagggctagc	4620
aaattctttg	gcctttcttc	ggtagcgatg	aacaaccggt	cattcaattt	ggcattggat	4680
actctttttg	cttcaacctgc	agagatttgc	gatctggtga	cgagggaaat	ggtaaagacc	4740
agtaacgata	ttgtgcataa	tattggatcg	aattccaaca	cggacgcgct	tcaaaagagc	4800
cttcaagttg	gtgcctcagc	agtagaaaaa	tacgacgagt	ctactctttc	tactaaagaa	4860
actgacgtat	attcccttgt	ttctgctctg	gctaaggaga	aatctcctct	atcttcttcc	4920
tcattctttgt	cgtctgaggg	acatctcacc	tttaaggaga	ttgataggac	atggaacacc	4980
cccgtctctt	tcgggtaccgc	taaaactaca	tcctattctg	tttctgaaga	cgctctcaat	5040
gtcctctctt	cagccgtggt	ggacttttaga	aggaatgttg	tggtatgctac	taaactctctg	5100
tacgaagttg	cagctgtttg	tagtgtcatg	agtaaagagg	aggatgtgctg	ttctctcagat	5160
agaaagatta	tgggaatggt	ggaacaagaa	tcgcccgtta	tgcaagacat	tggcattgac	5220
cgatttgcta	gtcttgttag	tacagttgct	acccccaaac	agcatcgag	attctttacag	5280
acagtaaacg	attacaaaaa	ttatctcatt	agaaaagttg	catcgaatcc	ccttctctct	5340
tcaagattgg	gaggaatatc	ccctactagt	ggtaacaccg	attacaacct	taaagctgta	5400
tatgatgggtg	ttgtttcttc	ctcatcatca	atgacccccct	cgtccatgct	tgtctctgac	5460
agattctggt	cgggagtatt	ttctcagttc	ctagagactg	gcccttcaat	gtttgccgat	5520
gctggctcatg	gaggtagtaa	catgttccaa	atcactgcac	ctaaacttta	cggttctaga	5580
gtcaacacct	acgcagctct	gagctctggc	gttgagcggc	tgagagactc	tatttcttct	5640
gcgactcagg	aaagaaagaa	taggattgca	aagagcatcg	aagctctgga	aacgttcgta	5700
accgatgtgg	tggggggaga	tactttggat	caattgctga	aggcccagaa	catgtacaac	5760

aaactgtcag	atattacttc	caactctatc	tatagtgatt	tcggaaacat	tgactgcgct	5820
aaaatcatga	agaatgtgac	gagcaagaaa	atgaccgcta	gacaacaatc	agatactatt	5880
cttagctctc	tcttgacaga	actcgctggc	ctggtacaca	aacaacaacc	tcaattggct	5940
actcaatttg	ctctggcgag	ccatgtttatc	aaggc aaagt	atgtcactaa	tgacctcaat	6000
aatatccacg	agaaggaaac	attcagttcaa	ttgatggccg	tgcccggtgt	tgccgattac	6060
tataatgtgt	cggcagctgc	catgtgtcag	cgtctagtgt	cttccgacgt	aacaatgttc	6120
ctcggcggaa	ccatgctcca	acaaggcctg	ttcgtttcat	tcctttctta	caacgtactt	6180
ttctcccagg	tttctgataa	tattaaaatg	aacgaattga	acgatgaaac	aaagtctctt	6240
ttggttaaac	tggtaggatt	ttgcggtaca	gtttcagatg	cgctaggatc	taggcacgtg	6300
tcttcaatta	gacgtgtaca	gaacgaagag	gataagaaat	tagacaggag	ttttgttaca	6360
tcactttatt	cagcatacag	agatttgagg	aagaagactg	aactatacac	ggaaactgat	6420
actattaaca	aacttttcgg	acatcaaaac	tttatgtctt	acgaatcttc	catgctcaag	6480
aggacttctc	tggtacatga	cgctgtttcc	ggccctaggc	caagaaggta	cagcaccctt	6540
gaggatgtac	ttgaggctcc	ttccacgggt	cacaaatcgt	tcatggtttc	ttaccacagag	6600
agggcagctg	cttctaggcg	agtgaagagg	gctggactca	gggctctggc	tgataacagg	6660
atggaatctc	tttacgggga	agaagtcttg	aacgatatga	ggtcttcggc	ggtctcttcc	6720
gaaatgatgg	atatagagta	tggtgaggga	ggattcatga	tgatgattag	tgatgatgag	6780
gatgatattg	cctttattga	ttccgaagaa	gagtctgaat	catctactga	tttctcttca	6840
tcagatgaat	attccgattc	atccgatgag	tatgattttg	atgatgataa	taatggccag	6900
tctccttatt	caactacatc	ttattcgtat	gatgctctag	accgtctgaa	ttctgccgct	6960
aagcctctta	ctgccatcta	cgggtgcagg	ggagaagggtg	aagacgatga	ggaaaatgac	7020
ctctatgaag	aagaacaaga	aaggaggaga	cgtcgctcat	caaagatggg	gaagatcctt	7080
agagatcttc	atgagagtga	tgatgacgac	gatgactact	ttgatgacga	atttgatggc	7140
gaacgttcaa	tgtcagaaac	tattgcaacc	agaagagctg	gccgtattca	atatggtcca	7200
ggtttcctat	ctcattctaa	tattcttaac	cgcccggtca	aagcacgcgc	tttcttgaca	7260
cgaggcaaga	aattcaggcc	ttctgcgtac	gatagattct	ttatggagga	tgacgattcc	7320
ctcctcttct	ctgacgaatc	taccacttct	tcttctctct	ccgattctcc	attctcctcc	7380
ttcagcaagg	ggagaaaatg	caagcgccga	acaagcgagg	accaatgtgc	ctttgttaag	7440
agagttgtac	gtgcttttgt	gcccaccaga	gtacaatga	tcaatggctg	agttagcatg	7500
atcacccag	tgactagtga	aaatacacga	ggattctatg	aaaattacca	gaaggccaac	7560
aagagggaag	gggcgcgtct	gatcgaagaa	tacaaaattg	ttaagggtgc	ttcagctacc	7620
ttgcccgacg	aatacgtaga	gggtagagca	tctaaacaag	tgtctcccag	ggaactgagg	7680
aggtctctta	tcaaggcagc	tgcttatgtt	gcccgcaccc	aagaaaagtaa	cttgaatatt	7740
atctttgacg	ctctcaccac	aacatcaaac	gccactctag	ttaatgaccc	atctactctt	7800
ttgggtgata	cacttttgtt	cgccaaacaa	ctagaggcaa	ttaccgagag	gaggaatagg	7860
ctaattgaaag	acctaactga	aatctctcct	tcacttttca	catcattcgg	tgatgcaagt	7920
aaagacaccc	aaatgcatggc	cgatggcaaa	cagatcggtt	caggaggaaa	tttcaagtct	7980
gccggctatc	taggtgtccc	tctcagaact	cttgcttcat	gtattaaggg	cactaataca	8040
gttgatcgtc	ttttggctac	caaaaataag	aaccatctcg	aatggatgac	cacagccgct	8100
attgtttttg	cacgttcatt	caacgatact	actttccatg	cactcgaaga	tacactaaaa	8160
atgacctcgg	ctttgacaga	catgtacagc	gctttcacca	acctcgctcg	atcggaaacat	8220
tctcagcgtc	taaaagttaa	gagtactctt	ttagattcta	ttttcaacac	taggatggct	8280
cacactgaag	cagtcatggg	tctcgtatac	cctacagcgt	tcatcaacca	tgaaatgccc	8340
tctgattaca	cacagcgtag	agagatgcaa	tcactcgctc	ttaacattct	taggggagtt	8400
aattgtagcc	aattgccacg	aaaggatatt	ggagacactg	ctggcctggt	gacctttatt	8460
acatcacgta	aattttgcagg	ttatggagga	gaaaggggag	gtttgtctct	atacagaatg	8520
tccattgttg	atgctctttc	ttgccccctc	gacaatcggc	tcaagggagc	agtcctctta	8580
gaggtaggaa	agtggcagga	tatgggagag	gaaatcttct	acaagaggag	caacgatctg	8640
gtcgattttt	gttcaaagaa	caatattttc	ctggaaaatg	ccgtagggtcc	tattgctagg	8700
tttgttccca	atggaactaa	catggctgat	attggcatga	ccgatatcat	ttctagaaca	8760
gtcaaggatg	acgcttcaat	gatcaggctt	aggcgcgtag	aagagggcgc	tgccgcagca	8820
ggaaaattca	ttacagcctc	agccatgggt	aatttgtacg	gaggtattga	taccgtttgtg	8880
aaactaactg	aaaaactata	cgaactcgtt	gttctgctcc	aagattcaga	ctcgttcaat	8940
acaccaacag	aaatggccac	tgctattatc	aaccgtatga	agtcgaggga	acataaggct	9000
ctcaaaacac	cattcggggg	agatattgcc	acctataaga	acttcccac	ctcctctgaa	9060
gcaattgtag	ttagagccaa	ggaaatgcgt	aactctatta	gcactatcgt	gatggacatt	9120
tccaagtcaa	ggggaatcaa	ctcgttctca	tcccgcagtg	gttctacttt	ggccaagatt	9180
ttcacatctg	aatttgaaaag	gatactagaa	acatctgctg	ttctttcaaa	tacaaaggcc	9240
aatctgagaa	ccattgagaa	tagactcgcc	gaacactaca	acaaactaaa	acaattcagc	9300
catattagta	atgatggact	ttccgagaca	cgcgcagctg	ttgccgtaat	tgctgaatct	9360
ttaacccccg	tgtatgcgga	tgacaccagc	gagagaggag	catctgttag	tgaactattg	9420



acagacaata	ctctcctcaa	atttatgtgt	caaaatgaac	tgaaaaaacat	tgaagaggca	9480
aaacgtcacg	tgaccgccgc	aattgaaggt	tcatcccaac	tgcacgaaaa	aatgttgagc	9540
ctgctcgttg	cctcagccga	catcaaccgt	atgtccgccc	aaaataacct	cgaatgtaag	9600
aaattgactg	aaggaatatg	taacttttga	ccaatgacta	acgaccaagc	tggtagattc	9660
ataaagcaca	aagaaaacag	tatctggctg	aagaccgatg	aagaaaaata	caccagtctc	9720
atcaaggaca	atgatcagcg	tagagttagt	aaaaccatcc	tgcgaattgt	agaggacaat	9780
agaaatgcaa	ccatccgttc	tctgtctacg	tctcttttgt	ttggaaaaata	tgccatgaac	9840
gacatttttg	cacttgatga	tgccgatatt	aagaatatgg	acaaactcat	tgaaaaacta	9900
ggcgaaagca	tgcgagagaa	ggcatctctc	tctagctcgg	ccattttctc	ttctctcatc	9960
tctaacacaa	catctctctc	tcttctctcc	agttctctcc	catcatcatc	atcttctctc	10020
ttctcaatgg	actattcaaa	caaccttgcc	aaaactatcc	cctacatgcc	tatcgtcttc	10080
caaaacaaac	aatctaatgt	caattcttct	gacgcatcat	cctcatcacc	atcatcatct	10140
tcttcttcat	ctgccaatat	tgataatgtt	gagcacaana	aagtggctct	ccaacaactt	10200
caaacacaa	aatctaacga	tttgagtaac	gtactttctg	ttaccacca	gcacagattt	10260
cgctctcata	atcaagctgc	aactgtttgc	atcttcaacg	gaaggcaaca	cgcagagaca	10320
gttgttgcta	taccaaagtc	aaacaaggct	aataataatg	ccaccgtttc	cgcaggccaa	10380
ggaattctta	cccgctttct	agcccttgaa	aatgtttcct	ccaccagcat	gcaattgcct	10440
ccatcatcat	catcatcatc	aaatggagat	gataataagg	taccagtaac	tgtcaggctt	10500
aaccagtagc	ccaactcaat	cttatcatct	attgaaaacg	catcagaatt	taaggacttg	10560
aaggaagcag	aaaggaana	cgatctggcc	atccaggcag	cttccaccac	agaaacaaag	10620
gaaatggtea	ccgtgtctaa	gtgcccctct	gctaaccaga	ctgccatcac	tgccatctct	10680
caagctaaat	cccttaagaa	aagtgccttc	gaattattgg	aaagagttat	caaggcagtc	10740
gaggtttaca	ccccagattc	atctattgca	gccgtttctc	ttcccgtaa	tgagagattc	10800
atggtttctt	cttctctggg	atcaggatct	gtcctctctt	catcatctct	ctctctaccc	10860
tctctctctt	ctctaatgt	gacagcatat	tcaactatg	cttacggaaa	attgaagaac	10920
attgatgaaa	atactgaaga	aggggcagaa	actgtccaga	aaaacatggt	cgaacaagat	10980
gctgccgttc	gcatecctct	tctagtatca	tatgtcccat	tcagcgaaat	gatgagacgt	11040
gctattgaca	agttgaacga	atactaccaa	ctgattgatg	ccatcaaaac	aaagatcgtg	11100
tcagacacta	aacaggcttc	ctcatggggc	atcaaggaaa	ggacacaagg	gcttgatatg	11160
gacaagaac	aggtgatctc	aaagattaat	aacttgcaac	aaaacttttc	aaacgaatca	11220
gacaagataa	agatggctat	tagtgttttg	gacaacaaaa	ggaacgaatt	agagcttcag	11280
aacaacaaaa	ctaggagctt	tattgaaact	acaaagagcc	gtatcgaggc	tggtggagga	11340
gatgtagcaa	acttcaagga	gattatcgat	tacgaaaaca	catctgaaaa	tgacaacaat	11400
ctcttcaga	gcttgaaagc	attcgctgct	gataactcgg	ggacagttta	cacccccact	11460
gacatgagca	atggaagaga	cacaaaatca	gacagtaaat	ttgtcgacat	tgacaacaaa	11520
cagattctcg	agggaggaat	caaaactcatc	aatgagggac	aaaataactgt	aaaggtagac	11580
ttttcaaagg	ctttggaggc	tttccctaga	caatccaacg	gtgcttcaga	gcctgtatct	11640
tcttcagttg	tggaagaggag	acagcgagaa	cgtctcaagg	ctgtcgagat	gtttatggca	11700
ataatgatgg	agcgcaccga	gagtcgtgag	aagaggttgg	cagattcggc	tgctcagttg	11760
aatactgtta	ataattgaga	agaactggtt	aatagtggta	tggttaacat	caagagtga	11820
aggtctcacg	agattagga	tcaagcacaa	atcgttgaaa	gcactgcact	aaactccatc	11880
aacgacgaga	ttgtagagtc	tctcttcacc	ctctctttgg	gagcacgagt	cgaccagctc	11940
ttgatcaagg	tagatagagt	aggaagtatc	caacagcagc	aacagcaaca	gcagcaacag	12000
cagcagcttc	ccaaattgac	agctacagaa	cagagaaaag	aacaacata	cgctgcagat	12060
agggttgttt	acgatccttc	atacacctgc	ttcctgcaac	ctcttcacga	gacaattaaa	12120
cgattttctt	ctgtctataa	ttcaaagaac	aagggtcctc	tcagtaacac	acgtgggtgt	12180
ccactagcgg	atgccgactt	gcaactgatg	accatcactg	acttgtctag	gtctgtactc	12240
gactcttctt	ccacttcttc	caagaaaatg	ctgtacgaaa	atgttccctc	atcaattgtt	12300
cctggactct	gccagcaatt	cgcaatgatg	atcaccaacg	agcacaaggc	cactcatact	12360
tctcctcatt	cattcaattg	cgagacaana	agatccctga	acgacgtgac	agaaatgttg	12420
aacgtctgca	cttcatccag	tgacggctct	ccgctgagac	acgatgtact	aaaatgttta	12480
gagtccaaca	atggttacgt	caaagatttt	ggattcactc	accgccaaaa	ggttgcggtg	12540
atcacccctg	ttaatacact	tctgggaggt	actttcagtg	gaaatgttgc	acctaatact	12600
gttatccttc	ctactttcta	gttgtttaac	tgccagagag	ttgaaaatga	caaatgttag	12660
tccatggtta	acaggacaac	cgacaagaat	gtggctgagc	cacccaagtc	atctgcaagc	12720
atcgtggaga	ctcttgctcg	cacgctctcc	aacgccgagc	acctttactt	cccttcaag	12780
gaccagaggc	gacacttcaa	ctccatcacc	gacgccatca	tttctgggat	gagcggcgaa	12840
tcttcatctc	aatt					

aacaaggaag	aaactctccg	tctaaacagg	ctttcagaca	ttaacaacgt	gagacattat	13140
ggcactgatg	ttcatgtggc	aggcgcaaac	tctgcatgga	gaattgggtga	ggtgggtgaga	13200
gccgcctcct	cattccctga	cggagataag	gaatcggcta	tgaaaaagat	gcttcttcta	13260
ggatctgtat	ctgccatctc	tgtcaaaaaa	tctgccagtc	acattaacga	tcctactgct	13320
ttgttgagca	ccaacacgtc	tatccagaat	ctggccaagg	aagctttccc	agaccctggt	13380
tgttctctcta	attacttggg	gtctgtcgaa	tctacgttcg	ccactcaact	cgcctaccgc	13440
cagcgctgt	tcctaaccgg	agatgacgaa	aacgttacaa	ccgtttcaaa	tatctgccct	13500
atggatttga	tgggaagtac	aaagcgctac	aatgacgctt	tcaacaacat	ctttggctct	13560
aaaatgacat	ctactaacia	aaaggggtca	aattgtgaaa	atctactgaa	atctgccatg	13620
tctaacgttc	ctgctatcaa	cactgccttt	ggagcctttg	aagaagcttc	atcttctgtc	13680
aggaataggc	tttctccctt	ttatgaagac	agcaccaaat	attcgtccaa	ccaacttgct	13740
gtacaggcca	tgaccgatac	tgtgtgggat	gctttgtccg	ctgtttctac	tgttgtcggg	13800
cgccagaatg	gcagaaatac	tcttctttct	ctccctactt	ctattacttc	tatcgcaacc	13860
agtggccgtc	catcactctc	ttattcttcg	gacatgaaat	ctaacctcat	caagacaatt	13920
tcccgcacat	atagagacgc	tagcctcctg	tcaatgggag	acagccaagt	agctgcaggt	13980
tcttctctct	ttactctttt	ccttcgttct	tcttccatcc	ctgtcaccac	cagccaggat	14040
ggaaatgttg	cagcagcgga	aattgtttct	gggactattc	tcgacaagac	tgtggagatc	14100
aacaagagat	tcgagatgct	tggaggagga	aaaatggctg	ccgggagtc	tgaagctcgt	14160
gccatccagc	gcaatacaat	gtcctctatt	ctccagatga	acgaaaatga	actcgcctcg	14220
gacttgtgcg	aaattgaaaa	taaaattgag	actaggcaac	tgagggatgc	tttccaggat	14280
ttgaagaggt	ctatgctgat	gactccagga	ggcgtgggag	ccatttcttc	tggagcaagt	14340
accaacaatg	ttcccttttc	tcttctcatg	tcacgtgtcg	atgcattccg	cggctctctg	14400
atgaacaaca	acagtgccaa	tgtaatggaa	gctgtggata	gtttcaatac	tactcctttg	14460
ctcgttaggc	acatgatggt	ggatagtggg	aagtcccccg	ttcccatggc	caaggaaatt	14520
aggagcatgc	tgaccaaac	aagagctctc	accgcccgcg	ctctactgag	cgaatcttcc	14580
cctcttctca	ctgaaatctg	cctctacaac	accgcgcgca	ctcaaccaga	aagggcagtc	14640
gacagactac	taacttcagc	ctatctagta	aaacaagcta	aaagattcga	cggagtgtac	14700
ccagccttcc	ctgccgccct	cacctgcgct	tctcacctca	tgttttcttc	catggattcc	14760
catacaaagt	catctttcat	ggacaacatc	aaattgcaca	tgactgatac	tcaatgcttc	14820
ttcaagaaca	ttgaacgatt	tgagaaattc	ttgggaagat	atggggacga	atacgccatg	14880
ttccacaagc	aaaattgtaa	ctgccccctc	catctccacc	acactttttac	tcctcagat	14940
aacgagcatc	tggatccctc	tttcgcatte	gcccgcgcag	aagtctccat	ggaagaaatt	15000
agagccacac	cctatcaggc	caacaagctt	attagtgcga	aacattacgt	gatgaacatg	15060
tccaagatcg	attctagagt	aacaggatct	tcctctctta	agaaggttag	cgaatggact	15120
gaaatgagaa	tgaactccaa	ctttaatgga	acatttgaac	catcaagact	cgcctctctc	15180
aactctggca	tgacaacggc	aggagtcaac	ctcgacgtta	ttgtcaaacc	aaataatgca	15240
agaagtgtac	taggaatatt	ggaatgtcat	gcgcagcacg	tgtgcaccgc	cgacgccaag	15300
ggaactgtcg	cttcagccat	gccagccgtc	ttccaggcaa	ccgatggaaa	cggtaacgaa	15360
tctgaactga	tccagaatgc	tctgccaaag	aacagataca	tccaaaagag	cacaatgaac	15420
gctcaaaactg	tcgtgtttgc	taatgttttg	gaacaactta	tcgccgatct	tggaaagggt	15480
atcgtgaacg	aactggccgg	caccatcgct	gaatctgtac	cagaaaagcg	atatgaaaac	15540
accaagggaaa	tgattgatag	actaggctct	gacgacctct	tcaaactctaa	taataatgga	15600
ggagtagaat	caatggatta	tgaagatagc	gaaacaacat	ccaacaatgg	tcccgtctct	15660
atctcagaag	ccatgaagaa	tgcctcttat	cacacactaa	tttccggcaa	ggcagctcgc	15720
ccggaaaatg	taccattcgc	ctcatgcgcc	agcggccctc	tcgcctttga	tttcttctct	15780
tcaaagggag	atacattcga	agaaaagaac	gccgaacaag	gtgcagcagc	tgccgtatcc	15840
tctacctatt	cttctctctc	taacactact	cttcgtaagc	atttggctcg	agttttcgaa	15900
gccatctcta	agcaagtaac	tgatgctgaa	ttcaaggata	tcctcaacga	tatcgaaact	15960
aatatttctt	ctgactatac	taactgtcca	ccaaatacta	acaaaaatgc	ctttgtctta	16020
gctatcaaga	gagaattcag	cagaattggt	tccttcttaa	ccattcttcg	taagaacatt	16080
acacccgcac	tagtcgaccc	taaaggcgcg	ttacacgaga	aagtagccat	ctatttgacc	16140
cttcttttcaa	ccaaatcaaa	actagaaaac	tttttccaat	acggctctcag	taattcgtcc	16200
tctagtgatc	ttagccatct	aaaaccattt	aattgtagca	acaatgtcaa	gaatatgaa	16260
gacacattca	tgtacagaaa	tgtccaccct	attcttatta	tggccctccc	agaaaaattc	16320
acagctctct	tgcaacagga	acaaatggac	cccgatactg	ccattgaaag	cagacgctcc	16380
cttaccacct	tccttaatac	tcccaacact	gcttcaatgg	cgaacggtgc	aagagccgct	16440
gtgggtgcag	gaggaggaaa	cccaatgggc	ttgtatcttt	cttcccacat	tcttcacgag	16500
tctaccgtca	acaatcaaaa	ccccgtcaca	gacaccacag	aaaacgtcaa	ctatcattcc	16560
tctgtttacac	aagatcctgt	tatggtatgt	aaccccttca	aggattctgc	taggttgatc	16620
gttaacaaca	acaatactgg	aattgatgtc	ttgaatgata	agtcgtgcaa	ctacttgcaa	16680
gtatccatgc	catctgaatc	atctggcctc	gtcaccaata	ctggatgctc	ttcttcttct	16740

```

tcctcatctt cgtctgatac cttcaagtac gtcaggagag acaatacgcc tgtgaatctt 16800
ccccgtgtca caccagccgt tctctgttct gatgcttctt ctaatctctt ggacgtgttc 16860
tccagggcag atattgtcct cgaaaacatg aacgtgagat ttggtttcat gcccgagatt 16920
attgctgccg tctccaaatt caaggggctg accaaggaag aggttattaa gcaaatgggt 16980
tctcagaaca acatcaacaa caacagcaac aacaacaacg gaaatgggaa gaaaacaacc 17040
gtcgtatccag tcaactggga tattgttatt accaatgcca cattccccga cactcgctct 17100
ctatacactg cagcaaatgg aggaacatca tcattcaaatt ggggagatat caacgacaga 17160
aaaatgcacg ccaaggcttt ccccaccttc tttattggta acccaaccgc cgccgcaaca 17220
gctaacggag tgccctcttac atctgaggga atttccctca ctgaagaaaa acgcaagaaa 17280
atcgcaggca tctctgaagg atcaattggc acgggggctc tgcgtgcagc cgccaacacc 17340
cgctctcat ccgacatgga acctgtcatg aagggatgga acaacattgt tcagcttcaa 17400
caaacattca agaaagcttc agataaactc actcatcttt tgagatcggg aggaattcca 17460
cccagaagcc aagaaacaaa cgctattatt aacaagatgc acgacagctt caagacattg 17520
gaggaatgtc gtaggggtgat ccaagacgag gctgctctgc tcgttgccac cagcgatctt 17580
ttgaccggtg gctacggagg agatgctgct ctggccatgg tttctccagt acgtccagaa 17640
atgactggtc ttattgggtc aatctccgag ccagtttagag gtattagcca cttgttgaaa 17700
ctgggaggtg tttctgctgc taacgcagct atccgcaagc gcctcaacct acctacatcc 17760
aacgggaaaa cactaccaga acatggaatc gtacacaaat cagccaagac acttttgctt 17820
gattcagact ctattagcaa cctatacaac actgatcttc aagacgttgt ctctaacgct 17880
agggataaca acaatttggg aagaattatg caatctttgg gacttaaggg gaataatgca 17940
ggggatttgg tttattctgc tagacaactg accggacctt ttactgtacc agaattgga 18000
aacaatcgcg atcttaccaa gcgtcaagct atccttaaaa tgctcatttc taaccctgaa 18060
attctagaaa atgttgacaga taccatttac cttacaacag gtaaaaaatgc tctcgcaccg 18120
gtatctgctc aggaaatggc ttgtgctgct ctgacagtcg gaggaagtgg aggaggaaaa 18180
ctgtcatcag acgacaatgt tcaatccctg aaccgccttt attttcgggt ctgag 18234

```

<210> 61  
 <211> 6025  
 <212> PRT  
 <213> SHRIMP

```

<400> 61
Met Asp Gln Tyr Pro Glu Val Arg Asp Thr Pro Gln Thr Glu Gln Glu
1      5      10      15
Gln Ala Ala Ala Gln Gln Gln Gln Ala Ala Thr Thr Thr Ala Ala Ala
20     25     30
Ala Ala Ala Ala Ala Pro Thr Gln Tyr Ser Asn Thr Val Ser Ala Glu
35     40     45
Thr Leu Ser Ala Ile Ser Glu Asp Gly Lys Leu Glu Arg Ser Ile Ala
50     55     60
Ala Ser Cys Trp Ile Asn Asn Leu Asn Pro Asp Glu Lys Met Ala Gln
65     70     75     80
Arg Val Gln Phe His Pro Leu Ser Ser Thr Thr Thr Tyr Asp Ser Glu
85     90     95
Asn Val Asn Pro Gly Ser Ser Val Val Phe Leu Lys Pro Arg Ala Leu
100    105    110
Pro Thr Gly Gly Thr Cys Leu Ala Pro Asn Tyr Ile Ala Val Pro Thr
115    120    125
Leu Arg Ala Ala Ser Glu Ile Ile Asp Ser Ile Ala Ser Thr Ser Leu
130    135    140
Tyr Gln Cys Ser Met Phe Asn Ser Trp Asn Leu Ile Pro Ile Phe Met
145    150    155    160
Ser Asn Ser Lys His Ser Gln Phe Gly Asp Arg Val Ile Lys Arg Ser
165    170    175
Met Ile Arg Asn Cys Phe Ser Lys Gln Lys Asn Val Glu Asn Leu Leu
180    185    190
Lys Glu Leu Arg Arg Arg Lys Val Asn Ala Ala Lys Ala Phe Ser His
195    200    205
Ala Val Gln Gln Lys Ser Ala Val Asn Thr Ala Ala Trp Asn Ala Gly
210    215    220
Ser Ala Ala Asn Leu Glu Lys Leu Val Asp Phe Cys Lys Leu Lys Tyr

```

WO 01/38351

144

PCT/US00/28888

225					230					235				240
Ser	Pro	Asp	Arg	Lys	Tyr	Lys	Ala	Gly	Gly	Leu	Phe	Ser	Ala	Ser
				245					250					255
Thr	Ala	Gln	Ser	Gln	Ser	Gly	Thr	Ser	Ser	Ser	Val	Glu	His	Thr
			260					265					270	
Ser	Asn	Asp	Phe	Leu	Leu	Asp	Ile	Leu	Lys	Arg	His	Lys	Gly	Thr
		275				280						285		
Leu	Asp	Leu	Asp	Ser	Ala	Thr	Asn	Thr	Phe	Asp	Thr	Ala	Leu	Ser
	290					295					300			
Val	Phe	Thr	Glu	Phe	Lys	Glu	Gln	Ala	Arg	Ala	Ala	Val	Asp	Ala
305					310					315				320
Ala	Asp	Ser	Asp	His	Leu	Ser	Ala	Ser	Asp	Pro	Ile	Phe	Ser	Ile
				325					330					335
Arg	His	Asn	Ser	Arg	Arg	Glu	Gly	Ile	Leu	Asp	Ser	Val	Pro	Asn
			340					345					350	
Gly	Met	Leu	Ala	Pro	Arg	Ser	Lys	Tyr	Ser	Val	Ala	Glu	Tyr	Leu
		355					360					365		
Ala	Asp	Arg	Asp	Glu	Ser	Ala	Asp	Ile	Ala	Ala	Lys	Ile	Gly	Thr
	370					375					380			
Ile	Ala	Thr	Asp	Phe	Glu	Ala	Leu	Arg	Gly	Asp	Asn	Asn	Lys	Arg
385					390				395					400
Ala	Asp	Thr	Ser	Val	Asp	Asp	Leu	Lys	Glu	Ser	Leu	Ala	Asp	Ser
				405					410					415
Glu	Lys	Thr	Ser	Ile	Lys	Asn	Thr	Gly	Asp	Ile	Asn	Ser	Val	Thr
			420					425					430	
Ile	Pro	Thr	Asp	Thr	Glu	Glu	Tyr	Glu	Phe	Ser	Leu	His	Ile	Thr
	435						440					445		
Leu	Phe	Ala	Gln	Ala	Phe	Leu	Glu	Thr	Met	Gly	Ser	Leu	Leu	Ser
	450					455				460				
Ala	Phe	Gly	Val	Gln	Phe	Pro	Phe	Ser	Asp	Glu	Gly	Phe	Ala	Ala
465					470					475				480
Glu	Arg	Ile	Ile	Arg	Lys	Thr	Asp	Pro	Asp	Thr	Gly	Lys	Val	Ser
				485					490					495
Met	Asp	Pro	Ser	Ser	Leu	Ser	Asp	Gln	Tyr	Leu	Leu	Leu	Val	Gly
			500					505					510	
Phe	Gln	Val	Ser	Pro	Phe	His	Val	Ser	Asp	Pro	Lys	Asp	Ile	Val
	515						520						525	
Gly	Arg	Gln	Val	Thr	Pro	Asn	Thr	Pro	Ile	Leu	Ser	Ile	Ile	Thr
	530					535					540			
Ser	Lys	Asn	Asp	Lys	Asn	Glu	Thr	Ser	Thr	Ile	Ile	Asn	Phe	Arg
545					550					555				560
Arg	Leu	Leu	Val	Asn	Asp	Thr	Val	Leu	Arg	Asp	Ala	Thr	Gln	Asn
				565					570					575
Ser	Thr	Ser	Thr	Pro	Ser	Gln	Arg	Arg	Val	Pro	Thr	Ala	Ala	Gly
				580				585					590	
Pro	Lys	Lys	Pro	Met	Leu	Ser	Gly	Cys	Leu	Pro	Ile	Ile	Arg	Gly
		595					600					605		
Gln	Val	Val	Thr	Arg	Glu	Ser	Asp	Asp	Met	Ile	Ser	Gly	Leu	Val
	610					615					620			
Asp	Trp	Tyr	Ile	Ser	Leu	Gly	Val	Tyr	Tyr	Ala	Met	Gly	Ser	Ser
625					630					635				640
Ala	Ala	Ile	Ala	Ala	Gly	His	Gln	Arg	Ala	Ser	Ala	Glu	Ser	Ile
				645					650					655
Ser	Pro	Met	Met	Lys	Lys	Phe	Ser	Lys	Lys	Gly	Gly	Lys	Tyr	Thr
			660					665					670	
Glu	Glu	Lys	Arg	Ile	Lys	Lys	Ala	Met	Arg	Arg	Asn	Ala	Asp	Arg
		675					680					685		
Ala	Arg	Ile	Leu	Ala	Leu	Leu	Gly	Gln	Thr	Asp	Ala	Gln	Tyr	Gly
	690					695				700				
Val	Glu	His	Asn	Ser	Thr	Leu	Asp	Ser	Phe	Trp	Ser	Ser	Asn	Ala
705					710					715				720

WO 01/38351

145

PCT/US00/28888

Ile	Arg	Ala	Lys	Ala	Lys	Glu	Asp	Ala	Leu	Ser	Arg	Ala	Glu	Ile	Leu
				725					730					735	
Ala	Val	Arg	Lys	Gln	Leu	Asp	Gly	Lys	Cys	Ser	Ser	Ser	Arg	Asp	Glu
			740					745					750		
Tyr	Ser	Met	Val	Glu	Arg	Tyr	Leu	Arg	Asp	Ser	Phe	Phe	Arg	Ser	Val
		755					760					765			
Asn	Arg	Ser	Gly	Gly	Gly	Tyr	Glu	Met	Phe	Asp	Gln	Gly	Phe	Asp	Met
	770					775					780				
Gly	Arg	Phe	Ala	Asp	Phe	Leu	Ser	Asp	Asn	Ser	Ala	Ala	Arg	Asn	Ala
785					790					795					800
Trp	Gln	Gln	Tyr	Ala	Glu	Val	Met	Arg	Gly	Leu	Ser	Lys	His	Glu	Lys
				805					810					815	
Arg	Val	Phe	Asn	Ile	Glu	Gly	Leu	Phe	Ser	Ala	Leu	Asn	Ser	Phe	Lys
			820					825					830		
Phe	Pro	Leu	Val	Pro	Glu	Gln	Gly	Arg	Lys	Lys	Thr	Val	Gly	Gly	Arg
		835					840					845			
His	Arg	Leu	Asn	Asn	Leu	Lys	Ala	Ala	Asn	Lys	Ile	Ile	Asn	Gly	Ile
	850					855					860				
Thr	Glu	Met	Thr	Leu	Gln	Ser	Ala	Ile	Asp	Gly	Thr	Gly	Ile	Ser	Asp
865					870					875					880
Ile	Ile	Gly	Ser	Val	Ser	Asp	Gly	Trp	Gly	Asn	Thr	Thr	Ala	Gln	Pro
			885						890					895	
Ser	Arg	Val	Lys	Ala	Leu	Lys	Thr	Leu	Ser	Asn	Phe	Ser	Gly	Asn	Gly
			900					905					910		
Asn	Val	Val	Ser	Ile	Pro	Val	Ser	Arg	Ala	Val	Lys	Cys	Ala	Ala	Gly
		915					920					925			
Ser	Arg	Gly	Gly	Glu	Thr	Leu	Lys	Cys	Val	Asp	Ile	Pro	Ser	Val	Ile
	930					935					940				
Ile	Ala	Asn	Leu	Ile	Ser	Asp	Lys	Arg	Ile	Leu	Asp	Gln	Leu	Cys	Gly
945					950					955					960
Gly	Gly	Met	Asn	Leu	Ala	His	Glu	Ile	Thr	Asn	Phe	Ile	Glu	Thr	Ile
			965						970					975	
Ala	Gly	Lys	Glu	His	Thr	Gly	Lys	Glu	Ser	Val	Phe	Leu	Ser	Pro	Arg
			980					985					990		
Leu	Ser	Val	Ile	Leu	Leu	Arg	Tyr	Ile	Trp	Phe	Asn	Ala	Ala	Val	Val
		995					1000					1005			
Ser	Leu	Thr	Asp	Ser	Asn	Ile	Lys	Met	Pro	Leu	Asn	Thr	Met	Ser	Glu
	1010					1015						1020			
Gly	Thr	Gly	Asp	Asp	Ile	Tyr	Arg	Asp	Tyr	Leu	Ala	Ile	Arg	Gly	Met
1025					1030					1035					1040
Val	Asn	Asn	Tyr	Asn	Ser	Ser	Leu	Ser	Ser	Ile	Ser	Val	Lys	Ala	Ile
			1045							1050				1055	
Ser	Asp	Arg	Tyr	Asn	Cys	Gly	Ser	Gly	Asn	Thr	Ser	Thr	Ser	Asn	Lys
		1060						1065					1070		
Asn	Val	Thr	Ile	Lys	Thr	Gln	Gly	Glu	Leu	Leu	Thr	Val	Leu	Gln	Gln
		1075					1080						1085		
Thr	Ala	Asn	Ala	Leu	Ser	Ala	Phe	Thr	Asn	Lys	Gly	Gly	Val	Gly	Ala
	1090					1095						1100			
Thr	Pro	Asp	Ala	Ala	Asn	Met	Ala	Asn	Val	Ile	Ser	Pro	Ile	Ala	Asn
1105					1110					1115					1120
Ala	Asp	Val	Val	Lys	Asn	Thr	Asn	Val	Val	Val	Ser	Gly	Leu	Asp	Arg
			1125							1130				1135	
Ile	Thr	Glu	Thr	Ile	Asn	Phe	Phe	Ser	Phe	Leu	Ser	Gln	Ile	Lys	Thr
		1140						1145					1150		
Met	Asn	Glu	Asn	Ile	Glu	Glu	Tyr	Leu	Arg	Arg	Tyr	Arg	Leu	Gly	Glu
	1155						1160					1165			
Gly	Leu	Asp	Lys	Lys	Glu	Leu	Asp	Asn	Phe	Val	Tyr	Pro	Asn	Ile	Ala
	1170					1175						1180			
Ala	Ile	Val	Lys	Arg	Glu	Leu	Gly	Val	Ser	Gly	Ser	Ala	Leu	Ser	Ser
1185					1190					1195					1200
Asn	Leu	Asp	Thr	Asp	Arg	Pro	Ile	Thr	Ile	Asp	Leu	Asn	Thr	Glu	Gln

				1205						1210						1215		
Pro	Leu	Ile	Val	Lys	Ala	Ser	Lys	Gly	Tyr	Asn	Arg	Tyr	Ala	Lys	Leu			
			1220						1225						1230			
Phe	Asn	Lys	Thr	Thr	Arg	Thr	Ala	Ala	Glu	Gln	Ala	Gln	Met	Glu	Gln			
		1235						1240						1245				
Tyr	Asn	Ala	Gln	Met	Ala	Ala	Asn	Thr	Ile	Pro	Gln	Leu	Val	Asn	Arg			
	1250						1255						1260					
Leu	Thr	Ile	Pro	Gly	Ser	Ile	Thr	Ala	Asp	Thr	Ala	Ile	Asn	Val	Val			
1265				1270						1275						1280		
Lys	Ala	Phe	Thr	Glu	Asn	Gly	Glu	Phe	Ser	Asn	Ala	Glu	Thr	His	Leu			
		1285						1290						1295				
Gly	Val	Met	Gly	Asn	Ala	Ile	Asn	Glu	Met	Gln	Pro	Leu	Phe	Thr	Asp			
		1300						1305						1310				
Gly	Phe	Asn	Val	Ala	Asn	Lys	Arg	Leu	Thr	Val	Asn	Val	Gly	Ser	Val			
	1315						1320						1325					
Ser	Lys	Leu	Ile	Gln	Asn	Gly	Leu	Thr	Val	Ser	Leu	Ile	Leu	Ala	His			
	1330						1335						1340					
Ser	Lys	Asp	Tyr	Val	Phe	Lys	Pro	Leu	Val	Gln	Asp	Phe	Ala	Lys	Leu			
1345				1350						1355						1360		
Leu	Leu	Ala	Val	Thr	Ala	Glu	Thr	Ser	Leu	Val	Val	Ser	Arg	Ser	Gln			
		1365						1370						1375				
Lys	Ser	Phe	Phe	Pro	Ile	Pro	Pro	Ser	Val	Phe	Ser	Ser	Gly	Gly	Leu			
		1380						1385						1390				
Phe	Lys	Ile	Asp	Arg	Glu	Met	Phe	Asp	Asn	Met	Lys	Thr	Asp	Tyr	Val			
	1395						1400						1405					
Val	Glu	Val	Ile	Arg	Gln	Leu	Ser	Lys	Asn	Ala	Thr	Ala	Ala	Ile	Glu			
	1410						1415						1420					
Arg	Cys	Asn	Asp	Ser	Asp	Ser	Ala	Ala	Arg	Ile	Ala	Lys	Ser	Gly	Glu			
1425				1430						1435						1440		
Ile	Tyr	Asn	Lys	Asp	Val	Ala	Ser	Thr	Thr	Ala	Ala	Pro	Gly	Thr	Ser			
		1445						1450						1455				
Ser	Ser	Ala	Leu	Thr	Leu	Phe	Ala	Asn	Asn	Leu	Gln	Asn	Pro	Ala	Lys			
		1460						1465						1470				
Val	Trp	Ser	Met	Gly	Ala	Leu	Pro	His	Phe	Asp	Met	Ala	Val	Val	Pro			
	1475						1480						1485					
Lys	Leu	His	Gly	Ile	Ser	His	Asp	Gln	Met	Phe	Arg	Leu	Ser	Thr	Tyr			
	1490						1495						1500					
Tyr	Gln	Gly	Ile	His	Lys	Met	Glu	Leu	Asn	Ser	Asp	Cys	Lys	Pro	Glu			
1505				1510						1515						1520		
Glu	Trp	Asp	Asn	Ser	Leu	Pro	Gly	Asn	Arg	Ala	Ser	Lys	Phe	Phe	Gly			
		1525						1530						1535				
Leu	Ser	Ser	Val	Ser	Asp	Asn	Asn	Arg	Ser	Phe	Asn	Leu	Ala	Leu	Asp			
		1540						1545						1550				
Thr	Leu	Leu	Asp	Ala	Glu	Ile	Cys	Asp	Leu	Val	Thr	Arg	Glu	Met	Val			
	1555						1560						1565					
Lys	Thr	Ser	Asn	Asp	Ile	Val	His	Asn	Ile	Gly	Ser	Asn	Ser	Asn	Thr			

WO 01/38351

147

PCT/US00/28888

Val Met Ser Lys Glu Glu Asp Val Arg Ser Ser Ser Arg Lys Ile Met	1700	1705	1710
Gly Met Val Glu Gln Glu Ser Pro Val Met Gln Asp Ile Gly Ile Asp	1715	1720	1725
Arg Ile Ala Ser Leu Val Ser Thr Val Ala Thr Pro Lys Gln His Arg	1730	1735	1740
Arg Phe Leu Gln Thr Val Asn Asp Tyr Lys Asn Tyr Leu Ile Arg Lys	1745	1750	1755
Val Asn Pro Leu Leu Ser Ser Arg Leu Gly Gly Ile Ser Pro Thr Ser	1765	1770	1775
Gly Asn Thr Asp Tyr Asn Leu Lys Ala Val Tyr Asp Gly Val Val Ser	1780	1785	1790
Ser Ser Ser Ser Met Thr Pro Ser Ser Met Ser Val Ser Asp Arg Phe	1795	1800	1805
Trp Ser Gly Val Phe Ser Gln Cys Leu Glu Thr Gly Pro Ser Met Phe	1810	1815	1820
Ala Asp Ala Gly His Gly Gly Ser Asn Met Phe Gln Ile Thr Ala Pro	1825	1830	1835
Lys Leu Tyr Gly Ser Arg Val Asn Thr Tyr Ala Ala Leu Ser Ser Gly	1845	1850	1855
Val Glu Arg Leu Arg Asp Ser Ile Ser Ser Ala Thr Gln Glu Arg Lys	1860	1865	1870
Asn Arg Ile Ala Lys Ser Ile Glu Ala Leu Glu Thr Phe Val Thr Asp	1875	1880	1885
Val Val Gly Gly Asp Thr Leu Asp Gln Leu Arg Lys Ala Gln Asn Met	1890	1895	1900
Tyr Asn Lys Leu Ser Asp Ile Thr Ser Asn Ser Ile Tyr Ser Asp Phe	1905	1910	1915
Gly Asn Ile Asp Cys Ala Lys Ile Met Lys Asn Val Thr Ser Lys Lys	1925	1930	1935
Met Thr Ala Arg Gln Gln Ser Asp Thr Ile Leu Ser Ser Leu Leu His	1940	1945	1950
Glu Leu Ala Gly Leu Val His Lys Gln Gln Pro Gln Leu Ala Thr Gln	1955	1960	1965
Phe Ala Ser His Val Ile Lys Ala Lys Tyr Val Thr Asn Asp Leu Asn	1970	1975	1980
Asn Ile His Glu Lys Glu Thr Phe Ser Gln Leu Met Ala Val Ala Gly	1985	1990	1995
Val Ala Asp Tyr Tyr Asn Val Ser Ala Ala Ala Met Cys Gln Arg Leu	2005	2010	2015
Val Ala Ser Asp Val Thr Met Phe Leu Gly Gly Thr Met Leu Gln Gln	2020	2025	2030
Gly Leu Phe Val Ser Phe Leu Leu Asn Asn Val Leu Phe Ser Gln Val	2035	2040	2045
Ser Asp Asn Ile Lys Met Asn Glu Leu Asn Asp Glu Thr Lys Ser Leu	2050	2055	2060
Leu Val Lys Leu Val Gly Phe Cys Gly Thr Val Ser Asp Ala Leu Gly	2065	2070	2075
Ser Arg His Val Ser Ser Ile Arg Arg Val Gln Asn Glu Glu Asp Lys	2085	2090	2095
Lys Leu Asp Arg Ser Phe Val Thr Ser Lys Ala Tyr Arg Asp Leu Arg	2100	2105	2110
Lys Lys Thr Glu Leu Tyr Arg Glu Thr Asp Thr Ile Asn Lys Leu Phe	2115	2120	2125
Gly His Gln Asn Phe Met Ser Tyr Glu Ser Ser Met Leu Lys Arg Thr	2130	2135	2140
Ser Leu Val His Asp Ala Val Ser Gly Pro Arg Pro Arg Arg Tyr Ser	2145	2150	2155
Thr Leu Glu Asp Val Leu Glu Ala Pro Ser Thr Val His Lys Ser Phe	2165	2170	2175
Met Val Ser Tyr Pro Glu Arg Ala Ala Ala Ser Arg Arg Val Lys Arg			

[illegible]



Leu	Glu	Trp	Met	Thr	Thr	Ala	Ala	Ile	Val	Phe	Ala	Arg	Ser	Phe	Asn
		2675					2680					2685			
Asp	Thr	Thr	Phe	His	Ala	Leu	Glu	Asp	Thr	Leu	Lys	Met	Thr	Ser	Ala
		2690					2695					2700			
Leu	Thr	Asp	Met	Tyr	Ser	Ala	Phe	Thr	Asn	Leu	Val	Gly	Ser	Glu	His
2705						2710					2715				2720
Ser	Gln	Arg	Leu	Lys	Val	Lys	Ser	Thr	Leu	Leu	Asp	Ser	Ile	Phe	Asn
						2725					2730				2735
Thr	Arg	Met	Ala	His	Thr	Glu	Ala	Val	Met	Gly	Leu	Val	Tyr	Pro	Thr
						2740									2750
Ala	Phe	Ile	Asn	His	Glu	Met	Pro	Ser	Asp	Tyr	Thr	Gln	Arg	Arg	Glu
						2755									2760
Met	Gln	Ser	Leu	Ala	Leu	Asn	Ile	Leu	Arg	Gly	Val	Asn	Cys	Ser	Gln
						2770						2780			2775
Leu	Pro	Arg	Lys	Asp	Ile	Gly	Asp	Thr	Ala	Gly	Leu	Leu	Thr	Phe	Ile
2785						2790					2795				2800
Thr	Ser	Arg	Lys	Phe	Ala	Gly	Tyr	Gly	Gly	Glu	Arg	Gly	Gly	Leu	Ser
						2805					2810				2815
Leu	Tyr	Arg	Met	Ser	Ile	Val	Asp	Ala	Leu	Ser	Cys	Pro	Ser	Asp	Asn
						2820									2830
Arg	Leu	Lys	Gly	Ala	Val	Ser	Leu	Glu	Val	Gly	Lys	Trp	Gln	Asp	Met
						2835									2840
Gly	Glu	Glu	Ile	Phe	Tyr	Lys	Arg	Ser	Asn	Asp	Leu	Val	Asp	Phe	Cys
						2850						2860			2855
Ser	Lys	Asn	Asn	Ile	Ser	Leu	Glu	Asn	Ala	Val	Gly	Pro	Ile	Ala	Arg
2865						2870					2875				2880
Phe	Val	Pro	Asn	Gly	Thr	Asn	Met	Ala	Asp	Ile	Gly	Met	Thr	Asp	Ile
						2885					2890				2895
Ile	Ser	Arg	Thr	Val	Lys	Asp	Asp	Ala	Ser	Met	Ile	Arg	Leu	Arg	Arg
						2900									2910
Ala	Glu	Glu	Gly	Ala	Gly	Ala	Ala	Gly	Lys	Phe	Ile	Thr	Ala	Ser	Ala
						2915									2920
Met	Gly	Asn	Leu	Tyr	Gly	Gly	Ile	Asp	Thr	Val	Val	Asn	Leu	Thr	Glu
						2930						2940			2935
Lys	Leu	Tyr	Asp	Ser	Phe	Val	Leu	Leu	Gln	Asp	Ser	Asp	Ser	Phe	Asn
2945						2950					2955				2960
Thr	Pro	Thr	Glu	Met	Ala	Thr	Ala	Ile	Ile	Asn	Arg	Met	Lys	Ser	Arg
						2965					2970				2975
Lys	His	Lys	Ala	Leu	Lys	Thr	Pro	Phe	Gly	Gly	Asp	Ile	Ala	Thr	Tyr
						2980									2990
Lys	Asn	Phe	Pro	Ser	Ser	Ser	Glu	Ala	Ile	Val	Val	Arg	Ala	Lys	Glu
						2995									3000
Met	Arg	Asn	Ser	Ile	Ser	Thr	Ile	Val	Met	Asp	Ile	Ser	Lys	Ser	Arg
						3010						3020			

WO 01/38351

150

PCT/US00/28888

3155	3160	3165
Asn Asn Leu Glu Cys Lys Lys Leu Thr Glu Gly Asn Ser Asn Phe Val		
3170	3175	3180
Pro Met Thr Asn Asp Gln Gly Gly Thr Phe Ile Lys His Lys Glu Thr		
3185	3190	3195
Gly Ile Trp Leu Lys Thr Asp Glu Glu Asn Asn Thr Ser Ser Ile Lys		3200
	3205	3210
Asp Asn Asp Gln Arg Arg Val Ala Lys Thr Ile Leu Ala Ile Val Glu		3215
	3220	3225
Asp Asn Arg Asn Ala Thr Ile Arg Ser Arg Leu Gln Ser Leu Cys Phe		3230
	3235	3240
Gly Lys Tyr Ala Met Asn Asp Ile Phe Ala Leu Asp Asp Ala Asp Ile		3245
	3250	3255
Lys Asn Met Asp Lys Leu Ile Glu Lys Leu Gly Glu Ala Glu Lys Asp		3260
3265	3270	3275
Ser Ser Ser Ala Ile Ser Ser Ser Ser Ser Ser Asn Thr Thr Ser Ser		3280
	3285	3290
Ser Ser Ser Pro Ser Ser Ser Pro Ser Ser Ser Ser Ser Ser Phe Ser		3295
	3300	3305
Met Asp Tyr Ser Asn Asn Leu Ala Lys Thr Ile Pro Tyr Met Pro Ile		3310
	3315	3320
Val Phe Gln Asn Lys Gln Ser Asn Val Asn Ser Ser Asp Ala Ser Ser		3325
	3330	3335
Ser Ser Pro Ser Ser Ser Ser Ser Ser Ser Ala Asn Ile Asp Asn Val		3340
3345	3350	3355
Glu His Lys Lys Val Gln Gln Leu Gln Thr Gln Glu Ser Asn Asp Leu		3360
	3365	3370
Ser Asn Val Leu Ser Val Thr Thr Lys His Arg Phe Ala Ser His Asn		3375
	3380	3385
Gln Ala Ala Thr Val Gly Ile Phe Asn Gly Arg Gln His Ala Glu Thr		3390
	3395	3400
Val Val Ala Ile Pro Asn Ala Asn Lys Ala Asn Asn Ala Thr Val		3405
	3410	3415
Ser Ala Gly Gln Gly Ile Leu Thr Arg Phe Ser Ala Pro Glu Asn Val		3420
3425	3430	3435
Ser Ser Thr Ser Met Gln Leu Pro Pro Ser Ser Ser Ser Ser Ser Asn		3440
	3445	3450
Gly Asp Asp Asn Lys Val Pro Val Thr Val Arg Leu Asn Gln Tyr Ala		3455
	3460	3465
Asn Ser Ile Leu Ser Ser Ile Glu Asn Ala Ser Glu Phe Lys Asp Leu		3470
	3475	3480
Lys Glu Ala Glu Arg Lys Ile Asp Leu Ala Ile Gln Ala Ala Ser Thr		3485
	3490	3495
Thr Glu Thr Lys Glu Met Val Thr Val Ser Lys Cys Pro Ser Ala Asn		3500
3505	3510	3515
Gln Thr Ala Ile Thr Ala Ile Ser Gln Ala Lys Ser Leu Lys Lys Ser		3520
	3525	3530
Ala Leu Glu Leu Leu Glu Arg Val Ile Lys Ala Val Glu Val Tyr Thr		3535
	3540	3545
Pro Asp Ser Ser Ile Ala Ala Val Ser Leu Pro Val Asn Gly Asp Ser		3550
	3555	3560
Met Val Ser Ser Ser Ser Gly Ser Gly Ser Ala Pro Ser Ser Ser Ser		3565
	3570	3575
Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Asn Val Thr Asp Tyr Phe Asn		3580
3585	3590	3595
Tyr Ala Tyr Gly Lys Leu Lys Asn Ile Asp Glu Asn Thr Glu Glu Gly		3600
	3605	3610
Ala Glu Thr Val Gln Lys Asn Met Val Glu Gln Asp Ala Ala Val Arg		3615
	3620	3625
Ile Pro Leu Leu Val Ser Tyr Ala Pro Phe Ser Glu Met Met Arg Arg		3630
	3635	3640
		3645

Ala Ile Asp Lys Leu Asn Glu Tyr Tyr Gln Leu Ile Asp Ala Ile Lys  
3650 3655 3660  
Thr Lys Ile Val Ser Asp Thr Lys Gln Ala Ser Ser Trp Ala Ile Lys  
3665 3670 3675 3680  
Glu Thr Asp Lys Glu Leu Asp Met Asp Lys Glu Gln Val Ile Ser Lys  
3685 3690 3695  
Ile Asn Asn Leu Gln Gln Asn Phe Ser Asn Glu Ser Asp Lys Ile Lys  
3700 3705 3710  
Met Ala Ile Ser Val Leu Asp Asn Lys Arg Asn Glu Leu Glu Leu Gln  
3715 3720 3725  
Asn Asn Lys Thr Arg Ser Phe Ile Glu Thr Thr Lys Ser Arg Ile Glu  
3730 3735 3740  
Ala Gly Gly Gly Asp Val Ala Asn Phe Lys Glu Ile Ile Asp Tyr Glu  
3745 3750 3755 3760  
Asn Thr Ser Glu Asn Asp Asn Asn Leu Phe Gln Ser Leu Lys Ala Phe  
3765 3770 3775  
Ala Ala Asp Asn Ser Gly Thr Val Tyr Thr Pro Thr Asp Met Ser Asn  
3780 3785 3790  
Gly Arg Asp Thr Lys Ser Asp Ser Lys Phe Val Asp Met Tyr Asn Lys  
3795 3800 3805  
Gln Ile Gly Gly Ile Lys Leu Ile Asn Glu Gly Gln Asn Thr Val Lys  
3810 3815 3820  
Val Asp Phe Ser Lys Ala Leu Glu Ala Phe Pro Arg Gln Ser Asn Gly  
3825 3830 3835 3840  
Ala Ser Glu Pro Val Ser Ser Ser Val Val Glu Arg Arg Gln Arg Glu  
3845 3850 3855  
Arg Leu Gln Ala Val Glu Met Phe Met Ala Ile Met Met Glu Arg Thr  
3860 3865 3870  
Glu Ser Leu Arg Lys Arg Leu Ala Asp Ser Ala Ala Gln Trp Asn Thr  
3875 3880 3885  
Val Asn Asn Val Glu Glu Thr Val Asn Ser Gly Met Val Asn Ile Lys  
3890 3895 3900  
Ser Leu Thr Glu Ile Arg Asn Gln Ala Gln Ile Ala Glu Ser Thr Ala  
3905 3910 3915 3920  
Leu Asn Ser Ile Asn Asp Glu Ile Val Glu Ser Pro Leu Thr Leu Ser  
3925 3930 3935  
Leu Gly Ala Arg Val Asp Gln Leu Leu Ile Lys Val Asp Arg Val Gly  
3940 3945 3950  
Ser Ile Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Leu Pro  
3955 3960 3965  
Lys Leu Thr Ala Thr Glu Gln Arg Lys Glu Gln Gln Tyr Ala Ala Asp  
3970 3975 3980  
Arg Val Val Tyr Asp Pro Ser Tyr Thr Cys Phe Leu Gln Pro Leu His  
3985 3990 3995 4000  
Glu Thr Ile Lys Arg Ile Ser Ser Val Tyr Asn Ser Lys Asn Lys Gly  
4005 4010 4015  
Pro Leu Ser Asn Thr Arg Gly Val Pro Thr Ser Asp Ala Asp Leu Gln  
4020 4025 4030  
Leu Met Thr Ile Thr Asp Leu Ser Arg Ser Val Leu Asp Ser Ser Ser  
4035 4040 4045  
Thr Ser Ser Lys Lys Met Leu Tyr Glu Asn Val Pro Ser Ser Ile Val  
4050 4055 4060  
Pro Gly Leu Cys Gln Gln Cys Ala Met Met Ile Thr Asn Val His Glu  
4065 4070 4075 4080  
Ala Thr His Thr Ser Pro His Ser Phe Asn Phe Glu Asn Lys Arg Ser  
4085 4090 4095  
Leu Lys Gln Leu Thr Glu Met Leu Asn Ala Ala Thr Ser Ser Ser Asp  
4100 4105 4110  
Gly Pro Ala Val Arg His Asp Val Leu Thr Met Leu Glu Ser Asn Asn  
4115 4120 4125  
Gly Tyr Val Lys Asp Phe Gly Phe Thr Gln Lys Val Ala Cys Ile Thr

WO 01/38351

152

PCT/US00/28888

4130	4135	4140
Pro Val Asn Thr Leu Leu Gly Gly Thr Phe Ser Gly Asn Val Ala Pro		
4145	4150	4155
Asn Thr Val Ile Leu Pro Thr Ser Glu Leu Phe Asn Cys Pro Gly Val		4160
	4165	4170
Glu Asn Asp Lys Phe Arg Ser Met Val Asn Arg Thr Thr Asp Lys Asn		4175
	4180	4185
Val Ala Asp Ala Pro Lys Ser Ser Ala Ser Ile Val Glu Thr Leu Ala		4190
	4195	4200
Arg Thr Ser Pro Asn Ala Glu His Leu Tyr Phe Pro Phe Lys Asp Gln		4205
	4210	4215
Arg Arg His Phe Asn Ser Ile Thr Asp Ala Ile Ile Ser Gly Met Ser		4220
4225	4230	4235
Gly Glu Ser Ser Ser Gln Leu Asn Thr Thr Cys Asp Gln Asn Leu Val		4240
	4245	4250
Asn Ile Asp Gln Thr Thr Gly Phe Pro Val Phe Thr Gly Arg Lys Gln		4255
	4260	4265
Gly Glu Arg Arg Ile Val His Thr Glu Asn Thr Met Glu Gly Ala Arg		4270
	4275	4280
Lys Asp Lys Asn Ser Gly Ile Pro Ser Cys Thr Lys Asp Arg Gln Thr		4285
	4290	4295
Tyr Ile Asp Met Gly Thr Lys Phe Met Val Ala Pro Gly Ser Leu Leu		4300
4305	4310	4315
Asn Ala Asn Lys Glu Glu Thr Leu Arg Leu Asn Arg Leu Ser Asp Ile		4320
	4325	4330
Asn Asn Val Arg His Tyr Gly Thr Asp Val His Val Ala Gly Ala Asn		4335
	4340	4345
Ser Ala Trp Arg Ile Gly Glu Val Val Arg Ala Ala Ser Ser Phe Pro		4350
	4355	4360
Asp Gly Asp Lys Glu Ser Ala Met Lys Lys Met Leu Leu Leu Gly Ser		4365
	4370	4375
Val Ser Ala Ile Ser Ala Gln Lys Ser Ala Ser His Ile Asn Asp Pro		4380
4385	4390	4395
Thr Ala Leu Leu Ser Thr Asn Thr Ser Ile Gln Asn Leu Val Lys Glu		4400
	4405	4410
Ala Phe Pro Asp Pro Val Cys Ser Ser Asn Tyr Leu Gly Ser Ala Glu		4415
	4420	4425
Ser Thr Phe Ala Thr Gln Leu Ala Tyr Arg Gln Arg Leu Phe Pro Asn		4430
	4435	4440
Gly Asp Asp Glu Asn Val Thr Thr Val Ser Asn Ile Cys Pro Met Asp		4445
	4450	4455
Leu Met Gly Ser Thr Lys Arg Tyr Asn Asp Ala Phe Asn Asn Ile Phe		4460
4465	4470	4475
Gly Ser Lys Met Thr Ser Thr Asn Lys Lys Gly Ser Asn Cys Glu Asn		4480
	4485	4490
Leu Leu Lys Ser Ala Met Ser Asn Val Pro Ala Ile Asn Thr Ala Phe		4495
	4500	4505
Gly Ala Phe Glu Glu Ala Ser Ser Ser Val Arg Asn Arg Leu Ser Pro		4510
	4515	4520
Leu Tyr Glu Asp Ser Thr Lys Tyr Ser Ser Asn Gln Leu Ala Val Gln		4525
	4530	4535
Ala Met Thr Asp Thr Ala Val Asp Ala Leu Ser Ala Val Ser Thr Val		4540
4545	4550	4555
Val Gly Arg Gln Asn Gly Arg Asn Thr Leu Leu Ser Leu Pro Thr Ser		4560
	4565	4570
Ile Thr Ser Ile Ala Thr Ser Gly Arg Pro Ser Leu Ser Tyr Ser Ser		4575
	4580	4585
Asp Met Lys Ser Asn Leu Ile Lys Thr Ile Ser Arg Ile Asn Arg Asp		4590
	4595	4600
Ala Ser Leu Leu Ser Met Gly Asp Ser Gln Val Ala Ala Gly Ser Ser		4605
4610	4615	4620

WO 01/38351

153

PCT/US00/28888

Phe	Phe	Asn	Ser	Phe	Leu	Arg	Ser	Ser	Ser	Ile	Pro	Val	Thr	Thr	Ser	4625	4630	4635	4640
Gln	Asp	Gly	Asn	Val	Ala	Ala	Ala	Glu	Ile	Val	Leu	Gly	Thr	Ile	Leu	4645	4650	4655	
Asp	Lys	Thr	Val	Glu	Ile	Asn	Lys	Arg	Phe	Glu	Met	Leu	Gly	Gly	Gly	4660	4665	4670	
Lys	Met	Val	Ala	Gly	Ser	Pro	Glu	Ala	Arg	Ala	Ile	Gln	Arg	Asn	Thr	4675	4680	4685	
Met	Ser	Ser	Ile	Leu	Gln	Met	Asn	Glu	Asn	Glu	Leu	Ala	Arg	Asp	Leu	4690	4695	4700	
Cys	Glu	Ile	Glu	Asn	Lys	Ile	Glu	Thr	Arg	Gln	Leu	Arg	Asp	Ala	Phe	4705	4710	4715	4720
Gln	Asp	Leu	Lys	Arg	Ser	Met	Leu	Met	Thr	Pro	Gly	Gly	Val	Gly	Ala	4725	4730	4735	
Ile	Ser	Ser	Gly	Ala	Ser	Thr	Asn	Asn	Val	Pro	Leu	Ser	Leu	Leu	Met	4740	4745	4750	
Ser	Arg	Val	Asp	Ala	Ser	Ser	Gly	Leu	Leu	Met	Asn	Asn	Asn	Ser	Ala	4755	4760	4765	
Asn	Val	Met	Glu	Ala	Val	Asp	Ser	Phe	Asn	Thr	Thr	Pro	Leu	Leu	Val	4770	4775	4780	
Arg	His	Met	Met	Leu	Asp	Ser	Gly	Lys	Ser	Pro	Val	Pro	Met	Ala	Lys	4785	4790	4795	4800
Glu	Ile	Arg	Ser	Met	Leu	Thr	Gln	Pro	Arg	Ala	Leu	Thr	Ala	Arg	Ala	4805	4810	4815	
Leu	Leu	Ser	Glu	Ser	Ser	Pro	Leu	Leu	Thr	Glu	Ile	Cys	Leu	Tyr	Asn	4820	4825	4830	
Thr	Arg	Asp	Thr	Gln	Pro	Glu	Arg	Ala	Val	Asp	Arg	Leu	Leu	Thr	Ser	4835	4840	4845	
Ala	Tyr	Leu	Val	Lys	Gln	Ala	Lys	Arg	Phe	Asp	Gly	Val	Asp	Pro	Ala	4850	4855	4860	
Phe	Pro	Ala	Ala	Leu	Thr	Cys	Ala	Ser	His	Leu	Met	Leu	Ser	Ser	Met	4865	4870	4875	4880
Asp	Ser	His	Thr	Lys	Ser	Ser	Phe	Met	Asp	Asn	Ile	Lys	Leu	His	Met	4885	4890	4895	
Thr	Asp	Thr	Gln	Cys	Phe	Phe	Lys	Asn	Ile	Glu	Arg	Phe	Glu	Lys	Phe	4900	4905	4910	
Leu	Gly	Arg	Tyr	Gly	Asp	Glu	Tyr	Ala	Met	Ser	His	Lys	Gln	Asn	Cys	4915	4920	4925	
Asn	Cys	Pro	Phe	His	Leu	His	His	Thr	Phe	Thr	Pro	Ser	Asp	Asn	Glu	4930	4935	4940	
His	Leu	Val	Ser	Ser	Phe	Ala	Phe	Ala	Arg	Pro	Glu	Val	Ser	Met	Glu	4945	4950	4955	4960
Glu	Ile	Arg	Ala	Thr	Pro	Tyr	Gln	Ala	Asn	Lys	Leu	Ile	Ser	Asp	Lys	4965	4970	4975	
His	Tyr	Val	Met	Asn	Met	Ser	Lys	Ile	Asp	Ser	Arg	Val	Thr	Gly	Ser	4980	4985	4990	
Ser	Leu	Leu	Lys	Lys	Val	Ser	Glu	Trp	Thr	Glu	Met	Arg	Met	Asn	Ser	4995	5000	5005	
Asn	Phe	Asn	Gly	Thr	Phe	Glu	Pro	Ser	Arg	Leu	Ala	Leu	Ser	Asn	Ser	5010	5015	5020	
Gly	Met	Thr	Thr	Ala	Gly	Val	Asn	Leu	Asp	Val	Ile	Val	Lys	Pro	Asn	5025	5030	5035	5040
Asn	Ala	Arg	Ser	Val	Leu	Gly	Ile	Cys	His	Arg	Gln	His	Val	Cys	Thr	5045	5050	5055	
Ala	Asp	Ala	Lys	Gly	Thr	Val	Ala	Ser	Ala	Met	Pro	Ala	Val	Phe	Gln	5060	5065	5070	
Ala	Thr	Asp	Gly	Asn	Gly	Asn	Glu	Ser	Glu	Leu	Ile	Gln	Asn	Ala	Leu	5075	5080	5085	
Pro	Arg	Asn	Arg	Tyr	Ile	Gln	Lys	Ser	Thr	Met	Asn	Ala	Gln	Thr	Val	5090	5095	5100	
Val	Phe	Ala	Asn	Val	Leu	Glu	Gln	Leu	Ile	Ala	Asp	Leu	Gly	Lys	Val				

```

5105          5110          5115          5120
Ile Val Asn Glu Leu Ala Gly Thr Ile Ala Glu Ser Val Pro Glu Ser
          5125          5130          5135
Val Tyr Glu Asn Thr Lys Glu Met Ile Asp Arg Leu Gly Ser Asp Asp
          5140          5145          5150
Leu Phe Lys Ser Asn Asn Asn Gly Gly Val Glu Ser Met Asp Tyr Glu
          5155          5160          5165
Asp Ser Glu Thr Thr Ser Asn Asn Gly Pro Val Leu Ile Ser Glu Ala
          5170          5175          5180
Met Lys Asn Ala Val Tyr His Thr Leu Ile Ser Gly Lys Ala Ala Arg
5185          5190          5195          5200
Pro Glu Asn Val Pro Phe Ala Ser Cys Ala Ser Gly Pro Leu Ala Phe
          5205          5210          5215
Asp Phe Leu Leu Ser Lys Gly Asp Thr Phe Glu Glu Lys Asn Ala Glu
          5220          5225          5230
Gln Gly Ala Ala Ala Val Ser Ser Thr Tyr Ser Ser Ser Ser Asn
          5235          5240          5245
Thr Thr Leu Arg Lys His Leu Ala Arg Val Phe Glu Ala Ile Ser Lys
          5250          5255          5260
Gln Val Thr Asp Ala Glu Phe Lys Asp Ile Leu Asn Asp Ile Glu Arg
5265          5270          5275          5280
Asn Ile Ser Ser Asp Tyr Thr Asn Cys Pro Pro Asn Thr Asn Gln Asn
          5285          5290          5295
Ala Phe Ala Ile Lys Arg Glu Phe Ser Arg Ile Val Ser Phe Leu Thr
          5300          5305          5310
Ile Leu Arg Lys Asn Ile Thr Pro Ala Leu Val Asp Pro Lys Gly Ala
          5315          5320          5325
Leu His Glu Lys Val Ala Ile Tyr Leu Thr Leu Leu Ser Thr Lys Ser
          5330          5335          5340
Lys Leu Glu Asn Phe Phe Gln Tyr Gly Leu Ser Asn Ser Ser Ser Val
5345          5350          5355          5360
Asp Leu Ser His Leu Lys Pro Ile Asn Cys Ser Asn Asn Val Lys Asn
          5365          5370          5375
Ile Glu Asp Thr Phe Met Tyr Arg Asn Val His Pro Ile Leu Ile Met
          5380          5385          5390
Ala Leu Pro Glu Asn Phe Thr Ala Leu Leu Gln Gln Glu Gln Met Asp
          5395          5400          5405
Pro Asp Thr Ala Ile Glu Ser Arg Arg Ser Leu Thr Thr Phe Leu Asn
          5410          5415          5420
His Pro Asn Thr Ala Ser Met Ala Asn Gly Ala Arg Ala Ala Val Gly
5425          5430          5435          5440
Ala Gly Gly Gly Asn Pro Met Gly Leu Ser Ser His Ile Leu His Glu
          5445          5450          5455
Ser Thr Val Thr Thr Ser Asn Pro Val Thr Asp Thr Thr Glu Asn Val
          5460          5465          5470
Asn Tyr His Ser Ser Val Thr Gln Asp Pro Val Met Val Val Asn Pro
          5475          5480          5485
Phe Lys Asp Ser Ala Arg Leu Ile Val Asn Asn Asn Asn Thr Gly Ile
          5490          5495          5500
Asp Val Leu Asn Asp Lys Ser Cys Asn Tyr Leu Gln Val Ser Met Pro
5505          5510          5515          5520
Ser Glu Ser Ser Gly Leu Val Thr Asn Thr Gly Cys Ser Ser Ser
          5525          5530          5535
Ser Ser Ser Ser Ser Asp Thr Phe Lys Tyr Val Arg Arg Asp Asn Thr
          5540          5545          5550
Pro Val Asn Leu Pro Arg Val Thr Pro Ala Val Leu Cys Ser Asp Ala
          5555          5560          5565
Ser Ser Asn Leu Leu Asp Val Phe Ser Arg Ala Asp Ile Val Leu Glu
          5570          5575          5580
Asn Met Asn Val Arg Phe Gly Phe Met Pro Glu Ile Ile Ala Ala Val
5585          5590          5595          5600

```

```

Ser Lys Phe Lys Gly Leu Thr Lys Glu Glu Val Ile Lys Gln Met Val
      5605      5610      5615
Ser Gln Asn Asn Ile Asn Asn Asn Ser Asn Asn Asn Asn Gly Asn Gly
      5620      5625      5630
Lys Lys Thr Thr Val Asp Pro Val Thr Gly Asp Ile Val Ile Thr Asn
      5635      5640      5645
Ala Thr Phe Pro Asp Trp Leu Tyr Thr Ala Ala Asn Gly Gly Thr Ser
      5650      5655      5660
Ser Phe Lys Trp Gly Asp Ile Asn Asp Arg Lys Met His Ala Lys Ala
      5665      5670      5675      5680
Phe Pro Thr Phe Phe Ile Gly Asn Pro Thr Ala Ala Ala Thr Ala Asn
      5685      5690      5695
Gly Val Pro Leu Thr Ser Glu Gly Ile Ser Leu Thr Glu Glu Lys Arg
      5700      5705      5710
Lys Lys Ile Ala Gly Ile Ser Glu Gly Ser Ile Gly Thr Gly Ala Leu
      5715      5720      5725
Arg Ala Ala Ala Asn Thr Arg Leu Ser Ser Asp Met Glu Pro Val Met
      5730      5735      5740
Lys Gly Trp Asn Asn Ile Val Gln Leu Gln Gln Thr Phe Lys Lys Ala
      5745      5750      5755      5760
Ser Asp Lys Leu Thr His Leu Leu Arg Ser Gly Gly Ile Pro Pro Arg
      5765      5770      5775
Ser Gln Glu Thr Asn Ala Ile Ile Asn Lys Met His Asp Ser Phe Lys
      5780      5785      5790
Thr Leu Glu Glu Cys Arg Arg Val Ile Gln Asp Glu Ala Ala Leu Leu
      5795      5800      5805
Val Ala Thr Ser Asp Leu Leu Thr Gly Gly Tyr Gly Gly Asp Ala Ala
      5810      5815      5820
Met Val Ser Pro Val Arg Pro Glu Met Thr Gly Leu Ile Gly Ala Ile
      5825      5830      5835      5840
Ser Ala Pro Val Arg Gly Ile Ser His Leu Leu Lys Leu Gly Gly Val
      5845      5850      5855
Ser Ala Ala Asn Ala Ala Ile Arg Lys Arg Leu Asn Leu Pro Thr Ser
      5860      5865      5870
Asn Gly Lys Thr Leu Pro Glu His Gly Ile Val His Lys Ser Ala Lys
      5875      5880      5885
Thr Leu Leu Leu Asp Ser Asp Ser Ile Ser Asn Leu Tyr Asn Thr Asp
      5890      5895      5900
Leu Gln Asp Val Val Ser Asn Ala Arg Asp Asn Asn Asn Leu Gly Arg
      5905      5910      5915      5920
Ile Met Gln Ser Leu Gly Leu Lys Gly Asn Asn Ala Gly Asp Leu Val
      5925      5930      5935
Tyr Ser Ala Arg Gln Leu Thr Asp Leu Ile Thr Val Pro Glu Tyr Gly
      5940      5945      5950
Asn Asn Arg Asp Leu Thr Lys Arg Gln Ala Ile Leu Lys Met Leu Ile
      5955      5960      5965
Ser Asn Pro Glu Ile Asn Val Ala Asp Thr Ile Tyr Leu Thr Thr Gly
      5970      5975      5980
Lys Asn Ala Pro Val Ser Ala Gln Glu Met Ala Cys Ala Ser Leu Thr
      5985      5990      5995      6000
Val Gly Gly Ser Gly Gly Gly Lys Leu Ser Ser Asp Asp Asn Val Gln
      6005      6010      6015
Ser Leu Asn Arg Leu Tyr Phe Arg Val
      6020      6025

```

<210> 62  
 <211> 2190  
 <212> DNA  
 <213> SHRIMP

<400> 63															
Met	Asp	Lys	Val	Cys	Val	Ile	Ser	Asn	Thr	Arg	Glu	Arg	Thr	Phe	Lys
1				5					10					15	
Val	Pro	Ala	Asp	Leu	Leu	Cys	Val	Ala	Thr	Glu	Pro	Glu	Ile	Ser	Thr
			20					25					30		
Lys	Glu	Glu	Asp	Ala	Gly	Ile	Glu	Ile	Glu	Thr	Arg	Val	Val	Val	Phe
		35					40					45			
Ser	Arg	Cys	Val	Ser	Val	Gln	Glu	Leu	His	Thr	Ile	Asn	Pro	Asn	Asp
	50					55					60				
Glu	Gly	Phe	Ser	Val	Gln	Leu	Phe	Lys	Asp	Tyr	Leu	Lys	Leu	Gln	Ser
65					70					75					80
Ala	Gln	Gly	Lys	Lys	Pro	Ile	Gly	Ile	Gln	Ile	Lys	Ala	Gly	Glu	Asp
				85					90					95	
Leu	Glu	Arg	Arg	Leu	Ile	Ser	Gly	Gly	Thr	Ala	Tyr	Leu	Asp	Pro	Ala
			100					105					110		
Thr	His	Leu	Phe	Tyr	Leu	Asp	Phe	Ser	Leu	Tyr	Pro	Asn	Tyr	Ser	Ile
		115					120					125			



Phe	Asn	Asp	Ile	Ser	Ser	Arg	Leu	Lys	Ile	Ile	Asp	Glu	Asp	Thr	Tyr
130	135										140				
Asn	Gly	Val	Val	Phe	Ser	Asn	Ser	Glu	Glu	Lys	Glu	Lys	Asp	Ala	Leu
145				150						155					160
Val	Leu	Ile	Arg	Val	Thr	Phe	Ser	Thr	His	Glu	Lys	Ala	Ile	Glu	Ala
				165					170					175	
Ala	Ile	Lys	Lys	Ile	Met	Leu	Arg	Lys	Val	Phe	Phe	Lys	Asp	Gly	Asp
			180					185				190			
Leu	Asp	Phe	Gly	Tyr	Leu	Arg	Ile	Pro	Lys	Ser	Lys	Leu	Asp	Lys	Phe
		195					200					205			
Thr	Pro	Tyr	Phe	Arg	Ser	Gln	Tyr	Gly	Ile	Val	Asn	Val	Glu	Lys	Asn
	210					215					220				
Ile	Pro	Gly	Tyr	Ile	Trp	Gly	Glu	Ile	Met	Lys	Gln	Arg	Val	Arg	Cys
225					230					235					240
Ser	Arg	Trp	Tyr	Leu	Tyr	Asn	Thr	Asp	Ser	Glu	Trp	Glu	Tyr	Lys	Asn
				245					250					255	
Val	Ala	Glu	Glu	Arg	Val	Gly	Pro	Arg	Gln	Leu	Val	Lys	Lys	Tyr	Gly
			260					265				270			
Ala	Lys	Cys	Glu	Asn	Leu	Cys	Phe	Arg	Asp	Ile	Asp	Leu	Arg	Lys	Lys
		275					280					285			
Glu	Ala	Lys	Glu	Lys	Arg	Asp	Ile	Glu	Arg	Glu	Thr	Glu	Ser	Arg	Tyr
	290					295					300				
Val	Val	Val	Thr	Leu	Thr	His	Lys	His	Glu	Met	Pro	Glu	Asn	Met	Pro
305				310						315					320
Tyr	Phe	Gly	Pro	Lys	Cys	Ser	Val	Val	Arg	Leu	Asp	Glu	Thr	Arg	Ile
				325					330					335	
Leu	Leu	Cys	Phe	Val	Asp	Glu	Ile	Ser	Tyr	Asn	Asp	Glu	Asp	Val	Asp
			340					345				350			
Glu	Ile	Leu	Ser	Glu	Asn	Arg	Ser	Leu	Arg	Asn	Val	Ser	Ile	Arg	His
		355					360					365			
Lys	Glu	Asn	Val	Pro	Val	His	Thr	Leu	Leu	Lys	Lys	Gly	Val	Ser	Ile
	370					375					380				
His	Ala	Arg	Phe	Thr	Leu	Asn	Gly	Leu	Asp	Asp	Ala	Leu	Ile	Ile	Leu
385				390						395					400
Lys	Arg	Ile	Pro	Lys	Thr	Tyr	Phe	Glu	Asp	Glu	Glu	Leu	Gln	Ala	Ala
				405					410					415	
Cys	Ala	His	Val	Asn	Leu	Glu	Gln	Tyr	Glu	Trp	Leu	Cys	Ser	Asn	Asn
			420					425				430			
Arg	Gly	Asn	Lys	Val	Glu	His	Val	Lys	Ser	Arg	Val	Val	Thr	Arg	Ala
		435					440					445			
Val	Lys	Arg	Arg	Arg	Lys	Cys	Arg	His	Trp	Ile	Tyr	Phe	Asp	Lys	Asp
	450					455					460				
Thr	Leu	Asn	Leu	Asn	Tyr	Lys	Tyr	Phe	Asp	Lys	Lys	Val	Thr	Ala	Ser
465				470						475					480
Met	Ala	Ser	Lys												

<400> 65															
Met	Cys	Thr	Leu	Lys	Thr	Tyr	Lys	Met	Thr	Thr	Ser	Thr	Glu	Ile	Ser
1				5					10					15	
Lys	Asn	Leu	Ser	Asp	Val	Leu	Ser	Ile	Lys	Ala	Thr	Gly	Asp	Trp	Cys
			20					25					30		
Ser	Asn	Ile	Lys	Thr	Val	Phe	Ser	Pro	Phe	Thr	Glu	Gly	Lys	Gly	Asn
		35				40						45			
Leu	Pro	Asn	Ser	Leu	Pro	Phe	Thr	Arg	Ser	Pro	Asn	Thr	Thr	Cys	Gly
	50					55					60				
Ser	Arg	Glu	Ala	Ala	Asn	Ala	Thr	Glu	His	Phe	Ile	Thr	Val	Phe	Ala
65					70					75				80	
Lys	Asp	Lys	Tyr	Glu	Arg	Lys	Arg	Val	Lys	Arg	Thr	Ile	Gly	Phe	Thr
			85						90				95		
Leu	Asp	Asn	Thr	Lys	Glu	Leu	Thr	Pro	Asn	Arg	Tyr	Leu	Val	Ala	Asp
			100					105					110		

WO 01/38351

159

PCT/US00/28888

Val	Tyr	Ser	Trp	Gln	Glu	Glu	Lys	Met	Val	Phe	Glu	Gly	Phe	Cys	Val
	115						120					125			
Pro	Pro	Gly	Lys	Ser	Gly	Thr	Phe	Val	Arg	Tyr	Ser	Asn	Glu	Asp	Lys
	130					135					140				
Ser	Phe	Leu	Leu	Ala	Asp	Thr	Gly	Arg	Tyr	Met	Lys	Lys	Lys	Tyr	Asp
145					150					155					160
Asp	Pro	Glu	Asn	Lys	Thr	Ser	Ser	Gly	Gly	Asp	Asp	Asp	Asp	Asp	Asp
			165					170						175	
Asp	Asp	Asp	Asp	Asp	Asp	Asn	Asn	Asn	Val	Asp	Val	Tyr	Glu	Glu	Asn
			180					185					190		
Asp	Pro	Arg	Asn	Val	Phe	Glu	Val	Glu	Lys	Asp	Glu	Lys	Tyr	Ala	Cys
	195						200					205			
Thr	Phe	Ser	Ile	Leu	Val	Tyr	Arg	Ala	Met	Lys	Lys	Ser	Pro	Pro	Val
	210					215					220				
Cys	Arg	Gly	Leu	Leu	Val	Glu	Thr	Asp	Gly	Pro	Ser	Ser	His	Pro	Lys
225					230					235					240
Arg	Ala	Pro	Ser	Ala	Phe	Asn	Pro	Phe	Gly	Gly	Ser	Ser	Met	Leu	Asn
				245					250					255	
Gly	Tyr	Gly	Ala	Gly	Ala	Asp	Ala	Leu	Glu	Glu	Glu	Asp	Glu	Val	Asp
			260					265					270		
Gly	Val	Pro	Glu	Arg	Glu	Arg	Ile	Thr	Asn	Phe	Ala	Leu	Lys	Arg	Gly
	275						280					285			
Pro	Ala	Thr	Gly	Gln	Asn	Phe	Val	Ser	Val	Lys	Leu	Glu	His	Asp	Gly
	290					295					300				
Ser	Lys	Ala	Asp	Leu	Tyr	Asn	Val	Thr	Cys	Phe	Ser	Lys	Gln	Arg	Gly
305				310						315					320
Val															

<210> 66  
 <211> 1197  
 <212> DNA  
 <213> SHRIMP

<400> 66  
 atgcaactca ttcttttctca tcatctaacc atggctgggc gtgtagagct cgtcactgga 60  
 cccatgtttg cgggcaagtc tacctacctg aaaaacatat accaacaaga aaatggaggc 120  
 aataaacatt gcctgtttgt caaacactcc ctagaacta ggtacggttg tggactgga 180  
 acaatagtc ctcacgcccga agaagtgttg gaaggttgta ctacagtttc ttctatcaag 240  
 gaactaatca gtgtgttacc agaagtgttg gatgtgattc tcattgacga agggcaattc 300  
 ttacacggatt tgggtctagt caatagactg gctgacaagg ggaaaaggat tgtgattgca 360  
 gcacttgatg gaacttctga ccagcaaatg ttcagtccta ttcataagct attgccttat 420  
 acaaattcca ttgttaagct agcatctaaa tgtatgattt gtaaaattga taccaaagaa 480  
 gctcctttta ctgtaagggt ttgtaaatgac aatgataata atgttatatg ttagaggagg 540  
 gctgaaatgt acgctgctgc ctgcccgggac tgttacaaaa aaattaacaa gaaaaagaac 600  
 aaggggaaac ttgttgact tgaaggaggt gacagggtgc gtaagagtac ccaagccaaa 660  
 ctcttggtga ccaataaaaa ctgcctctt tatggaggag aatatatgtg ctttcccgac 720  
 aggagcagcc atacgggtaa actcatcaat gattatttaa ctaagaaaat tgaactagat 780  
 gatcatgcag ctacttggt attttctgca aatagatggg aagttttag taaaattaag 840  
 cagttgttag acgatggaat ccatgttgtg atggatagat attactactc ggggattggt 900  
 ttctctttag ctagaggagt ggataccgtt gagtgggtgc ctgctagcga tgagggactt 960  
 cctcagcccg atcttgatt gttgatgctt tttagatttg aaaagtgttc aaatagggat 1020  
 acttttggtg tcgaaagatt tgagacaaat tccattcaag aacgtgctag agccctattc 1080  
 ctagacctcg caaataagga cgaaaagaat gtatggatta aggtagacgc tcgcggcacc 1140  
 attgaggagg tgcaaaactaa aattataaat attgtatata atattgttga agaataa 1197

<210> 67  
 <211> 394  
 <212> PRT  
 <213> SHRIMP

&lt;400&gt; 67

```

Met Gln Leu Ile Leu Ser His His Leu Thr Met Ala Gly Arg Val Glu
 1      5      10      15
Leu Val Thr Gly Pro Met Phe Ala Gly Lys Ser Thr Tyr Leu Lys Asn
      20      25      30
Ile Tyr Gln Gln Glu Asn Gly Gly Asn Lys His Cys Leu Phe Val Lys
      35      40      45
His Ser Leu Glu Thr Arg Tyr Gly Cys Gly Thr Gly Thr Ile Val Thr
      50      55      60
His Ala Gly Glu Val Ile Glu Gly Cys Thr Thr Val Ser Ser Ile Lys
65      70      75      80
Glu Leu Ile Ser Val Leu Pro Glu Val Val Asp Val Ile Leu Ile Asp
      85      90      95
Glu Gly Gln Phe Thr Asp Leu Val Leu Val Asn Arg Leu Ala Asp
      100      105      110
Lys Gly Lys Arg Ile Val Ile Ala Ala Leu Asp Gly Thr Ser Asp Gln
      115      120      125
Gln Met Phe Ser Pro Ile His Lys Leu Leu Pro Tyr Thr Asn Ser Ile
130      135      140
Val Lys Leu Ala Ser Lys Cys Met Ile Cys Lys Ile Asp Thr Lys Glu
145      150      155      160
Ala Pro Phe Thr Val Arg Phe Gly Asn Asp Asn Asp Asn Asn Val Ile
      165      170      175
Cys Val Gly Gly Ala Glu Met Tyr Ala Ala Ala Cys Arg Asp Cys Tyr
      180      185      190
Lys Lys Ile Asn Lys Lys Lys Asn Lys Gly Lys Leu Val Val Leu Glu
      195      200      205
Gly Gly Asp Arg Cys Gly Lys Ser Thr Gln Ala Lys Leu Leu Leu Thr
210      215      220
Asn Lys Asn Ser Pro Leu Tyr Gly Gly Glu Tyr Met Cys Phe Pro Asp
225      230      235      240
Arg Ser Ser His Thr Gly Lys Leu Ile Asn Asp Tyr Leu Thr Lys Lys
      245      250      255
Ile Glu Leu Asp Asp His Ala Ala His Leu Leu Phe Ser Ala Asn Arg
      260      265      270
Trp Glu Val Cys Ser Lys Ile Lys Gln Leu Leu Asp Asp Gly Ile His
      275      280      285
Val Val Met Asp Arg Tyr Tyr Tyr Ser Gly Ile Val Phe Ser Leu Arg
      290      295      300
Val Asp Thr Val Glu Trp Cys Ser Ala Ser Asp Glu Gly Leu Pro Gln
305      310      315      320
Pro Asp Leu Val Leu Leu Met Leu Leu Asp Val Glu Lys Cys Ser Asn
      325      330      335
Arg Asp Thr Phe Gly Val Glu Arg Phe Glu Thr Asn Ser Ile Gln Glu
      340      345      350
Arg Ala Arg Ala Leu Phe Leu Asp Leu Ala Asn Lys Asp Glu Lys Asn
      355      360      365
Val Trp Ile Lys Val Asp Arg Thr Ile Glu Glu Val Gln Thr Lys Ile
      370      375      380
Ile Asn Ile Val Tyr Asn Ile Val Glu Glu
385      390

```

&lt;210&gt; 68

&lt;211&gt; 486

&lt;212&gt; DNA

&lt;213&gt; SHRIMP

&lt;400&gt; 68

atgttaccta gaaagacttt gcccgacact gaaaatggtt atttgtctt ggacgagtct 60

WO 01/38351

161

PCT/US00/28888

```

cttctggaga aggtgtacta tgataacaac aatgaactga ttgtaagagt tgggtgggatt 120
tatatgcaga tatgcaagtc aaaatacatc ttccatcacg atgatccaga gaggttctttt 180
tatagtgtgt tggaggatta tcaccccatc aaagagattg ttgaacgact agcagaagag 240
gatgggggat ttttaggacc gtggggagttt ttatcgcgca aacaagtgaa cctccaacac 300
gggtgctaca aagctctttt gtcattgcca gaggacaaat attgtaacct attattaccc 360
cagcaaatga aaaccaacct ggaaaaaatg gaagaaatac agcgtactag actcattcac 420
tctagaacgt acaatacacc ccagatagaa ttgtctgacc agctagatgg atgtgttata 480
tgtaa                                           486

```

<210> 69  
 <211> 161  
 <212> PRT  
 <213> SHRIMP

```

<400> 69
Met Leu Pro Arg Lys Thr Leu Pro Asp Thr Glu Asn Gly Tyr Phe Val
 1          5          10          15
Leu Asp Glu Ser Leu Leu Glu Lys Val Tyr Tyr Asp Asn Asn Asn Glu
          20          25          30
Leu Ile Val Arg Val Gly Gly Ile Tyr Met Gln Ile Cys Lys Ser Lys
          35          40          45
Tyr Ile Phe His His Asp Asp Pro Glu Arg Phe Phe Tyr Ser Val Leu
          50          55          60
Glu Asp Tyr His Pro Ile Lys Glu Ile Val Glu Arg Leu Ala Glu Glu
          65          70          75          80
Asp Gly Val Phe Leu Gly Pro Trp Glu Phe Leu Ser Arg Lys Gln Val
          85          90          95
Asn Leu Gln His Gly Cys Tyr Lys Ala Leu Leu Ser Leu Pro Glu Asp
          100          105          110
Lys Tyr Cys Asn Leu Leu Leu Pro Gln Gln Met Lys Thr Asn Leu Glu
          115          120          125
Lys Met Glu Glu Ile Gln Arg Thr Arg Leu Ile His Ser Arg Thr Tyr
          130          135          140
Asn Thr Pro Gln Ile Glu Leu Ser Asp Gln Leu Asp Gly Cys Val Ile
          145          150          155          160
Cys

```

<210> 70  
 <211> 1926  
 <212> DNA  
 <213> SHRIMP

```

<400> 70
atggttgctt caactccgtg tccaggccca ggaccagttc caaccaaga acttctttct 60
acaaactttc ttgaagctca caagcttgct gtggaacttc ttctcccgtc ctacagtagt 120
gatgtagttt attgtgactc tgagacgtac accaaaccta taccgatttt tgggaacaag 180
agtatagttt ctaccattgg agactatgtc ttatcaaacc ccaatgaaga tgtgagttac 240
caaatggttt ctccgctctt agaaaaattt cccttgctat tccactgcac ttataagacg 300
aatgaagaag ataaagggtat tcctctgtgg aagaagttgt acaacaaaag aaaattcaaa 360
ctcctcaact cattgttggt tcataacaac aagaactgga ctctgttccc agctatcccc 420
ttgacaggg agaatatatg tgatgcttca ggaaggagtg ttcttatgag tgaaataatg 480
tccacgtcaa cttttcagac aatttgcaaa aacaacacac attacttgtt tgatatgtta 540
aatatggaac gtggcaaaca aggaggaggt tttcttcact tctttgcatc taggaagaat 600
tcttttacta actttgaaaa tgaagaaatg gactctcatg tgctcagtaa catagcgaaa 660
ttcatatgca atgaaaagga aaaactagac tctttcatac ctgccaacgg aaaaatacca 720
tgccctgata aaactaatga tgaagggtac atcccgctgg aaatagcaat tatggaagac 780
aattaccctg cattgctata tctcgtttgt aggtatggag catcttgggc aaacacatac 840
ggggatcata atgaatctct caaagcgttt gcaataagaa atgatgcaaa agattgtctg 900
gaaattatag agtttataag tgatcactac agtttcaaca aaaatgtgac gaaggaagaa 960

```

```

tttggttaaag agaagactgt agaatgtggt ggatgtttat atgatattga agacgagaaa 1020
cggttggttaca aactcccatg tggacatttc atgcatacat tttgcttgct taataagtgt 1080
tctaaagcta acttttagatg tggttaaagt ttccaaacct ttgatgacac aattttttaga 1140
aaatgtcccc caactataca atggaaaatg ggtataaacc aaacgactaa ccataaggaa 1200
atggatttgt tcaatcgtgc atttgacaca tatttagatt ttatttgctc atataacgctc 1260
aaattagaca aaaaatcaaa acctaaacac aaacctgaaa acaaaaaggt ggaagaagaa 1320
ctagcaaaaa ggacagcaga aattgaagag gccataaaga aaaaggaaga agaactagca 1380
aaaaggacag cagaaattga agaggccata aagaaaaagg aagaagaact agcaaaaagg 1440
acagcagaaa ttgaagaggc catgaagaaa aaggaagaag aagaactctc aaaatataat 1500
aaaataattg aaaagggaaa aagacgactg aatgaagaat gtgtcaagct gagagatatt 1560
tcaactgcag ccataaacat gtacaaagag aaagtgaaga ttaatggtgt attactaaaa 1620
gattccgac aggagttggc tgaggcgaaa gagaggttga ggaaaatttt attgctagaa 1680
gaagaaacaa aacttgacag atttttggtt agaccgaaac gtagtagaaga acgtatatctc 1740
ctaactaaag atgatgaaac gttagccttc aagttagccc tagaaaagaa aacggaggac 1800
ataattgcga agaaaaacaa ccaaaaaggc agtgaaagaa gagatggaga atatactata 1860
acttctcata ttgagaaact acctcaatcc actgctttgg ctagtgtgtg tgtgttaaac 1920
gaataa

```

<210> 71  
 <211> 637  
 <212> PRT  
 <213> SHRIMP

<400> 71

Met	Val	Ala	Ser	Thr	Pro	Cys	Pro	Gly	Pro	Gly	Pro	Val	Pro	Thr	Gln
1				5					10					15	
Glu	Leu	Leu	Ser	Thr	Asn	Phe	Leu	Glu	Ala	His	Lys	Leu	Val	Val	Glu
			20					25					30		
Leu	Leu	Leu	Pro	Ser	Tyr	Ser	Ser	Asp	Val	Val	Tyr	Cys	Asp	Ser	Glu
		35					40					45			
Thr	Tyr	Thr	Lys	Pro	Ile	Pro	Ile	Phe	Gly	Asn	Lys	Ser	Ile	Val	Ser
		50				55					60				
Thr	Ile	Gly	Asp	Tyr	Val	Leu	Ser	Asn	Pro	Asn	Glu	Asp	Val	Ser	Tyr
65					70					75					80
Gln	Met	Val	Ser	Ser	Val	Leu	Glu	Lys	Phe	Pro	Leu	Leu	Phe	His	Cys
				85					90					95	
Thr	Tyr	Lys	Thr	Asn	Glu	Glu	Asp	Lys	Gly	Ile	Pro	Leu	Trp	Lys	Lys
			100					105					110		
Leu	Tyr	Asn	Lys	Arg	Lys	Phe	Lys	Leu	Leu	Asn	Ser	Leu	Leu	Val	His
		115					120					125			
Asn	Asn	Lys	Asn	Trp	Thr	Pro	Val	Pro	Ala	Ile	Pro	Phe	Asp	Arg	Glu
		130				135					140				
Asn	Ile	Cys	Asp	Ala	Ser	Gly	Arg	Ser	Val	Leu	Met	Ser	Glu	Ile	Met
145					150					155					160
Ser	Thr	Ser	Thr	Phe	Gln	Thr	Ile	Cys	Lys	Asn	Asn	Thr	His	Tyr	Leu
				165					170					175	
Phe	Asp	Met	Leu	Asn	Met	Glu	Arg	Gly	Lys	Gln	Gly	Gly	Ser	Phe	Leu
			180					185					190		
His	Phe	Phe	Ala	Ser	Arg	Lys	Asn	Ser	Phe	Thr	Asn	Phe	Glu	Asn	Glu
		195					200					205			
Glu	Met	Asp	Ser	His	Val	Leu	Ser	Asn	Ile	Ala	Lys	Phe	Ile	Cys	Asn
		210				215					220				
Glu	Lys	Glu	Lys	Leu	Asp	Ser	Phe	Ile	Pro	Ala	Asn	Gly	Lys	Ile	Pro
225					230					235					240
Cys	Pro	Asp	Lys	Thr	Asn	Asp	Glu	Gly	Tyr	Ile	Pro	Leu	Glu	Ile	Ala
				245					250					255	
Ile	Met	Glu	Asp	Asn	Tyr	Pro	Ala	Leu	Leu	Tyr	Leu	Val	Cys	Arg	Tyr
				260				265					270		
Gly	Ala	Ser	Trp	Ala	Asn	Thr	Tyr	Gly	Asp	His	Asn	Glu	Ser	Leu	Lys
		275				280						285			
Ala	Phe	Ala	Ile	Arg	Asn	Asp	Ala	Lys	Asp	Cys	Leu	Glu	Ile	Ile	Glu

```

      290              295              300
Phe Ile Ser Asp His Tyr Ser Phe Asn Lys Asn Val Thr Lys Glu Glu
305              310              315              320
Phe Val Lys Glu Lys Thr Val Glu Cys Val Gly Cys Leu Tyr Asp Ile
              325              330              335
Glu Asp Glu Lys Arg Cys Tyr Lys Leu Pro Cys Gly His Phe Met His
              340              345              350
Thr Phe Cys Leu Ser Asn Lys Cys Ser Lys Ala Asn Phe Arg Cys Val
              355              360              365
Lys Cys Phe Gln Thr Phe Asp Thr Ile Phe Arg Lys Cys Pro Pro
              370              375              380
Thr Ile Gln Trp Lys Met Gly Ile Asn Gln Thr Thr Asn His Lys Glu
385              390              395              400
Met Asp Leu Phe Asn Arg Ala Phe Asp Thr Tyr Leu Asp Phe Ile Cys
              405              410              415
Ser Tyr Asn Val Lys Leu Asp Lys Lys Ser Lys Pro Lys His Lys Pro
              420              425              430
Glu Asn Lys Lys Val Glu Glu Glu Leu Ala Lys Arg Thr Ala Glu Ile
              435              440              445
Glu Glu Ala Ile Lys Lys Lys Glu Glu Glu Leu Ala Lys Arg Thr Ala
              450              455              460
Glu Ile Glu Glu Ala Ile Lys Lys Lys Glu Glu Glu Leu Ala Lys Arg
465              470              475              480
Thr Ala Glu Ile Glu Glu Ala Met Lys Lys Lys Glu Glu Glu Glu Leu
              485              490              495
Ser Lys Tyr Asn Lys Ile Ile Glu Lys Gly Lys Arg Arg Leu Asn Glu
              500              505              510
Glu Cys Val Lys Leu Arg Asp Ile Ser Thr Ala Ala Ile Asn Met Tyr
              515              520              525
Lys Glu Lys Val Arg Ile Asn Gly Val Leu Leu Lys Asp Ser Asp Gln
              530              535              540
Glu Leu Ala Glu Ala Lys Glu Arg Leu Arg Lys Ile Leu Leu Leu Glu
545              550              555              560
Glu Glu Thr Lys Leu Asp Arg Phe Leu Phe Arg Pro Lys Arg Val Glu
              565              570              575
Glu Arg Ile Phe Leu Thr Lys Asp Asp Glu Thr Leu Ala Phe Lys Leu
              580              585              590
Ala Leu Glu Lys Lys Thr Glu Asp Ile Ile Ala Lys Lys Asn Asn Gln
              595              600              605
Lys Gly Ser Arg Asp Gly Glu Tyr Thr Ile Thr Ser His Ile Glu Lys
              610              615              620
Leu Pro Gln Ser Thr Ala Ser Val Cys Val Leu Asn Glu
625              630              635

```

<210> 72  
 <211> 780  
 <212> DNA  
 <213> SHRIMP

```

<400> 72
atgtctactt gttcgaattt gttgtcagta tttggtggag gagattggac aacaacattc 60
ccattcgacc tcgtccatac acgtcaagag tgtgataaaa agagagagca agactactca 120
tttttcatta ctgaaacgtg taaaggagag aatattggta tacattcgta tgaacacacg 180
tcaaagatta ttgacacggg taataatgat tctacctcaa tagagggaact agaagtactg 240
aatatataca aagctataaaa ccatttagaa aatatcctaa aactcaacaa aggagaaaaa 300
attatactga tggatgtaga aacaatgata ctggaaactc ataaaatttt aatgaaaggg 360
attcttccca agggtaaaaa tggaagtttc agtacatgcg tacgctttgc tgtaaataag 420
aacaatgaac ggcattacta ccctgtattht gaaacagaga aagaagcgtt caattctata 480
caaaatctag tagattatta taatgaaatt gtagctcaca ccaatgacca aattaaaata 540
ataaaagcgt gcgcatattht catgtacaac tttctaactc tccacccttt caatgatggt 600

```

WO 01/38351

164

PCT/US00/28888

aatggaagaa cagctagatt attgtatagt tttctattga aaggtaatgg tatcgtagct 660  
catttttcac ccataacaca ccctagggat caatttggtg atacttttagt gtattttaga 720  
gaacatggag atggacgacc tttattgtat gttttgctgg aatcaataaa aaataagtaa 780

<210> 73  
<211> 255  
<212> PRT  
<213> SHRIMP

<400> 73  
Met Ser Thr Cys Ser Asn Leu Leu Ser Val Phe Gly Gly Gly Asp Trp  
1 5 10 15  
Thr Thr Thr Phe Pro Phe Asp Leu Val His Thr Arg Gln Glu Cys Asp  
20 25 30  
Lys Lys Arg Glu Gln Asp Tyr Ser Phe Phe Ile Thr Glu Thr Cys Lys  
35 40 45  
Gly Glu Asn Ile Gly Ile His Ser Tyr Glu His Thr Ser Lys Ile Ile  
50 55 60  
Asp Thr Gly Asn Asn Asp Ser Thr Ser Ile Glu Leu Glu Val Leu  
65 70 75 80  
Asn Ile Tyr Lys Ala Ile Asn His Leu Glu Asn Ile Leu Lys Leu Asn  
85 90 95  
Lys Gly Glu Lys Ile Ile Leu Met Asp Val Glu Thr Met Ile Thr His  
100 105 110  
Lys Ile Leu Met Lys Gly Ile Leu Pro Lys Gly Lys Asn Gly Ser Phe  
115 120 125  
Ser Thr Cys Val Arg Phe Ala Val Asn Lys Asn Asn Glu Arg His Tyr  
130 135 140  
Tyr Pro Val Phe Glu Thr Glu Lys Glu Ala Phe Asn Ser Ile Gln Asn  
145 150 155 160  
Leu Val Asp Tyr Tyr Asn Glu Ile Val Ala His Thr Asn Asp Gln Ile  
165 170 175  
Lys Ile Ile Lys Ala Cys Ala Tyr Phe Met Tyr Asn Phe Leu Thr Leu  
180 185 190  
His Pro Phe Asn Asp Gly Asn Gly Arg Thr Ala Arg Leu Lys Phe Leu  
195 200 205  
Leu Lys Gly Asn Gly Ile Val Pro His Phe Ser Pro Ile Thr His Pro  
210 215 220  
Arg Asp Gln Phe Val Asp Thr Leu Val Tyr Phe Arg Glu His Gly Asp  
225 230 235 240  
Gly Arg Pro Leu Leu Tyr Val Leu Leu Glu Ser Ile Lys Asn Lys  
245 250 255

<210> 74  
<211> 480  
<212> DNA  
<213> SHRIMP

<400> 74  
atggaggacc taaaatccac tatcgagaga gtatatgaag aaagagtggga gaatctagaa 60  
caatggacaa atactgtaga ggaagaagaa aggactgtct cagcaatcga ttctgtcctg 120  
gaggaacaaa aaaggccct ggacgcatgg gaagcagcga taaaggaacg agaaaacgac 180  
ctcgagtaa aagaagggat atctgcactc gttttcaacg cagcagacgc caaaacacgt 240  
aaagaattga taaatacgtg gatagccgaa agggaaacgt cagaaaaaag aagaaggaa 300  
gcaacctcta ccaataatca actgaagaac cagatgtcat ctctagtcaa cacaacaaa 360  
acactcaaaag aaaagtacaa caaatattac agaagaagtg ccataactcaa catgcaatac 420  
atcaataaca aaagggatta tgaagcaagt caattttggg tgtatacaaa caatgcataa 480



WO 01/38351

165

PCT/US00/28888

<210> 75  
<211> 159  
<212> PRT  
<213> SHRIMP

<400> 75  
Met Glu Asp Leu Lys Ser Thr Ile Glu Arg Val Tyr Glu Glu Arg Val  
1 5 10 15  
Glu Asn Leu Glu Gln Trp Thr Asn Thr Val Glu Glu Glu Glu Arg Thr  
20 25 30  
Val Ser Ala Ile Asp Ser Val Leu Glu Glu Gln Lys Arg Ala Leu Asp  
35 40 45  
Ala Trp Glu Ala Ala Ile Lys Glu Arg Glu Asn Asp Leu Ala Val Lys  
50 55 60  
Glu Gly Ile Ser Ala Leu Val Phe Asn Ala Ala Asp Ala Lys Thr Arg  
65 70 75 80  
Lys Glu Leu Ile Asn Thr Trp Ile Ala Glu Arg Glu Thr Ser Glu Lys  
85 90 95  
Arg Arg Lys Glu Ala Thr Ser Thr Asn Asn Gln Leu Lys Asn Gln Met  
100 105 110  
Ser Ser Leu Val Asn Thr Thr Lys Thr Leu Lys Glu Lys Tyr Asn Lys  
115 120 125  
Tyr Tyr Arg Arg Ser Ala Ile Leu Asn Met Gln Tyr Ile Asn Asn Lys  
130 135 140  
Arg Asp Tyr Glu Ala Ser Gln Phe Trp Val Tyr Thr Asn Asn Ala  
145 150 155

<210> 76  
<211> 321  
<212> DNA  
<213> SHRIMP

<400> 76  
atgcataagt tttcaaataa attttatttt ataataaaaag gtgttttaaat tataattttt 60  
gtccctgatg ttgtgttttc tatattttta ctgcctcctc ttggggtaag acataaaaac 120  
ggtggcggcg gaaacgagga acagaagagc ggacccagcc agaagcatca tatccctggt 180  
cctgttctta tatttgctct catcatcggt atcggtggca gtgtcgtcat catcatcggt 240  
gtccttatca gtgtcaggat cgctgtcctt ctttgggtccc atccatacat tcatgacggc 300  
caggacgagg ataccaattg a 321

<210> 77  
<211> 106  
<212> PRT  
<213> SHRIMP

<400> 77  
Met His Lys Phe Ser Asn Lys Phe Tyr Phe Ile Ile Lys Gly Val Leu  
1 5 10 15  
Ile Ile Ile Phe Val Pro Asp Val Val Phe Ser Ile Phe Leu Leu Pro  
20 25 30  
Pro Leu Gly Val Arg His Lys Asn Gly Gly Gly Gly Asn Glu Glu Gln  
35 40 45  
Lys Ser Gly Pro Ser Gln Lys His His Ile Pro Gly Pro Val Leu Ile  
50 55 60  
Phe Val Leu Ile Ile Val Ile Val Gly Ser Val Ile Ile Ile Gly  
65 70 75 80  
Val Leu Ile Ser Val Arg Ile Ala Val Leu Leu Trp Ser His Pro Tyr  
85 90 95  
Ile His Asp Gly Gln Asp Glu Asp Thr Asn  
100 105

<210>	78
<211>	1635
<212>	DNA
<213>	SHRIMP

<400>	78						
atgttcccgtc	aattctgttc	actatatctt	cttcaacgcc	gggttaacga	caatcttaga	60	
tccacggcgt	ccgcttcggc	agcggcctca	ttaaaaggag	atggaactga	atttataacc	120	
ggagaaccac	cttctcataa	aatgagggga	ccttcttata	gcggttttag	acctgatccg	180	
tgcgaggacc	cagaaggggt	atatgttgat	attgtagtgt	ctattttgca	gacaaataat	240	
atacaggtaa	caaaagaatg	ggaattgttt	tccgataagt	tgagaaaaat	gggtccatgg	300	
attgatagga	gcggaattga	gaataatggc	gaaggagaag	aagatggaga	tgaaaatgaa	360	
gacgggggtg	gaaatggggg	aagaattgaa	gacagagaag	cacatcgacg	aaaaatgatg	420	
aagaaattgt	cctttgtttg	aagagaagat	ccagtcgctg	tagatttacc	cacgtggcga	480	
gaaaacagta	cagaatttgc	acgtcgttta	acactcaagg	aattgtgcga	tttaatagtt	540	
gaatgtggat	gcatcaaatc	aaaagaggaa	ctctttgact	tcttttttga	agaaccgtgg	600	
gagattaaag	agcgtgctga	cgttagggggt	atggcaaaca	ggagtaaatt	caccaaggaa	660	
tcattaattg	actggttttt	tgagttcgac	acatatagta	aatgtgtagt	attttttgaa	720	
gcagtcaact	ggtacttgaa	atctcaagcg	tctccaattt	cattggtact	agatgatata	780	
tattgtttgt	tcttttccct	cataagacgc	caaacccttt	taactagggc	aaaaaaccca	840	
tcttttaacg	tggtcttcac	cttttctccc	acgcccgaca	caaagctttt	ggctatcgac	900	
gagtgcgtcg	aacacttttt	aaaatcagac	attaatatata	gccagatggc	attaactgaa	960	
agggactgct	tcttccctct	tttaactgaa	atgccccgcc	aacaaaaaaa	agtaaacacc	1020	
ttcctggaca	caatgaagag	acctacctta	tctcttctac	cttccacctc	ctctctctct	1080	
tcttccaaca	acaagagaaa	gagaaatact	gccgctgcc	atattcttct	tccagtgtag	1140	
aggagtaact	tttctacagc	atccaaatac	aagagactga	aaactgatga	tggggaaaaa	1200	
gcatacgacct	gtattcttat	cgaagggtat	gcgaattgaa	aaataagccc	tataaggatt	1260	
atggtaagaa	aatcaactat	tattccgaaa	gtgtttaacc	atcttttggt	ccctgtcttt	1320	
gcctctaaag	acactgggtgc	gaatatctta	ttttttatca	aaatgaaatc	ctttgcaagt	1380	
gcattctttac	tcttccctgg	acttttttaga	caccccaaac	aatttctcaa	cgggccgtgc	1440	
aaatggatga	ctctagcaga	aaacaacatc	aacgacaaca	acataaactc	ttccacgatg	1500	
tggagttaca	cgctagcaga	ttattgtcct	ctgggctatt	acacccaaga	gagccctcaa	1560	
ccctatcaga	cgtacggcaa	ttttacttcg	actacaaaaca	agagactaca	aaacgtgcag	1620	
ccattatact	tttaa					1635	

```
<210> 79
<211> 540
<212> PRT
<213> SHRIMP
```

<400> 79																
Met	Phe	Arg	Gln	Phe	Cys	Ser	Leu	Tyr	Leu	Leu	Gln	Arg	Arg	Val	Asn	
1				5					10					15		
Asp	Asn	Leu	Arg	Ser	Thr	Ala	Ser	Ala	Ser	Ala	Ala	Ala	Ser	Leu	Lys	
			20					25					30			
Gly	Asp	Gly	Thr	Glu	Phe	Ile	Thr	Gly	Glu	Pro	Pro	Ser	His	Lys	Met	
		35					40					45				
Arg	Gly	Pro	Ser	Tyr	Ser	Val	Leu	Gly	Pro	Asp	Pro	Cys	Glu	Asp	Pro	
	50					55					60					
Glu	Arg	Val	Tyr	Val	Asp	Ile	Val	Val	Ser	Ile	Leu	Gln	Thr	Asn	Asn	
65					70					75					80	
Ile	Gln	Val	Thr	Lys	Glu	Trp	Glu	Leu	Phe	Ser	Asp	Lys	Leu	Arg	Lys	
				85					90					95		
Leu	Gly	Pro	Trp	Ile	Asp	Arg	Ser	Gly	Ile	Glu	Asn	Asn	Gly	Glu	Gly	
			100					105					110			
Glu	Glu	Asp	Gly	Asp	Glu	Asn	Glu	Asp	Gly	Gly	Gly	Asn	Gly	Gly	Arg	
		115					120					125				
Ile	Glu	Asp	Arg	Glu	Ala	His	Arg	Arg	Lys	Met	Met	Lys	Lys	Leu	Ser	
	130					135						140				

WO 01/38351

167

PCT/US00/28888

Phe	Val	Gly	Arg	Glu	Asp	Pro	Val	Ala	Val	Asp	Leu	Pro	Thr	Trp	Arg
145					150					155					160
Glu	Asn	Ser	Thr	Glu	Phe	Ala	Arg	Arg	Leu	Thr	Leu	Lys	Glu	Leu	Cys
				165					170						175
Asp	Leu	Ile	Val	Glu	Cys	Gly	Cys	Ile	Lys	Ser	Lys	Glu	Glu	Leu	Phe
			180					185					190		
Asp	Phe	Ile	Phe	Glu	Glu	Pro	Trp	Glu	Ile	Lys	Glu	Ala	Ala	Asp	Val
		195					200					205			
Arg	Gly	Met	Ala	Asn	Arg	Ser	Lys	Phe	Thr	Lys	Glu	Ser	Leu	Ile	Asp
	210					215					220				
Trp	Phe	Phe	Glu	Phe	Asp	Thr	Tyr	Ser	Lys	Cys	Val	Val	Phe	Phe	Glu
225					230					235					240
Ala	Val	Asn	Trp	Tyr	Leu	Lys	Ser	Gln	Asp	Ile	Ser	Leu	Val	Leu	Asp
				245					250						255
Asp	Ile	Tyr	Cys	Cys	Val	Phe	Ser	Tyr	Ile	Arg	Arg	Gln	Thr	Phe	Leu
			260					265					270		
Thr	Arg	Ala	Lys	Asn	Pro	Ser	Leu	Thr	Val	Ala	Ser	Ser	Phe	Ser	Pro
		275					280					285			
Thr	Pro	Asp	Thr	Lys	Leu	Leu	Ala	Ile	Asp	Glu	Cys	Val	Gln	His	Phe
	290					295					300				
Leu	Lys	Ser	Asp	Ile	Asn	Ile	Ser	Gln	Met	Ala	Leu	Thr	Glu	Arg	Asp
305					310					315					320
Cys	Phe	Phe	Pro	Leu	Leu	Thr	Glu	Met	Pro	Arg	Gln	Gln	Lys	Lys	Val
				325					330						335
Asn	Thr	Phe	Leu	Asp	Thr	Met	Lys	Arg	Pro	Thr	Leu	Ser	Leu	Leu	Pro
			340					345					350		
Ser	Thr	Ser	Ser	Ser	Ser	Ser	Ser	Asn	Asn	Lys	Arg	Lys	Arg	Asn	Thr
		355					360					365			
Ala	Ala	Ala	Asn	Ile	Leu	Leu	Pro	Val	Tyr	Arg	Ser	Asn	Phe	Ser	Thr
	370					375					380				
Asn	Asn	Lys	Arg	Leu	Lys	Thr	Asp	Asp	Gly	Glu	Asn	Ala	Ser	Ala	Cys
385					390					395					400
Ile	Leu	Ile	Glu	Gly	Tyr	Ala	Asn	Gly	Lys	Ile	Ser	Pro	Ile	Arg	Ile
				405					410						415
Met	Val	Arg	Lys	Ser	Thr	Ile	Ile	Pro	Glu	Val	Phe	Asn	His	Leu	Leu
			420					425					430		
Phe	Pro	Val	Phe	Ala	Ser	Lys	Asp	Thr	Gly	Ala	Asn	Ile	Leu	Phe	Phe
		435					440					445			
Ile	Lys	Met	Lys	Ser	Phe	Ala	Ser	Ala	Ser	Leu	Leu	Leu	Pro	Gly	Leu
	450					455					460				
Phe	Arg	His	Pro	Lys	Gln	Phe	Leu	Asn	Gly	Pro	Cys	Lys	Trp	Met	Thr
465					470					475					480
Leu	Ala	Glu	Asn	Asn	Ile	Asn	Asp	Asn	Asn	Ile	Asn	Ser	Ser	Thr	Met
				485					490						495
Trp	Ser	Tyr	Thr	Leu	Ala	Asp	Tyr	Cys	Pro	Leu	Gly	Tyr	Tyr	Thr	Gln
			500					505					510		
Glu	Ser	Pro	Gln	Pro	Tyr	Gln	Thr	Cys	Gly	Asn	Phe	Thr	Ser	Thr	Thr
		515					520					525			
Asn	Lys	Arg	Leu	Gln	Asn	Val	Gln	Pro	Leu	Tyr	Phe				
	530					535					540				

<210> 80  
 <211> 582  
 <212> DNA  
 <213> SHRIMP

<400> 80  
 atggatgact ctacgagaaa acaacatcaa cgacaacaac ataaactctt ccacgatgtg 60  
 gagttacacg ctacgagatt attgtcctct gggctattac acccaagaga gccctcaacc 120  
 ctatcagaca tgcggcaatt ttacttcgac tacaacaag agactacaaa acgtgcagcc 180

```

attatacttt taaacactct tttggaatac tacaggacac cttcagaaga gtgggaaatt 240
ccgtttaatc tcttgcttaa tgtgatgaat aacaagtgga gtacactcat tccagggtgtc 300
aaaataagtg caggtatcat atcgaaactc ccatggacca tgaaaacaat gtacgagatt 360
gtttcttcgc ccaataataa taataacaac ggagactact attctacatg caggcgaatg 420
gtaatggaat atcctatcgg ggggtttattg cacacgcctg ccataactaa taagtatcca 480
cgctccagaa tggtcacctg tacaaagggc aaagaccacc agaagctata tgacatctct 540
agacaaatgt ttgatataat agaagcaaat ggacaactct ga 582

```

&lt;210&gt; 81

&lt;211&gt; 193

&lt;212&gt; PRT

&lt;213&gt; SHRIMP

&lt;400&gt; 81

```

Met Asp Asp Ser Ser Arg Lys Gln His Gln Arg Gln Gln His Lys Leu
1          5          10          15
Phe His Asp Val Glu Leu His Ala Ser Arg Leu Leu Ser Ser Gly Leu
20          25          30
Leu His Pro Arg Glu Pro Ser Thr Leu Ser Asp Met Arg Gln Phe Tyr
35          40          45
Phe Asp Tyr Lys Gln Glu Thr Thr Lys Arg Ala Ala Ile Ile Leu Leu
50          55          60
Asn Thr Leu Leu Glu Tyr Tyr Arg Thr Pro Ser Glu Glu Trp Glu Ile
65          70          75          80
Pro Phe Asn Leu Leu Asn Val Met Asn Asn Lys Trp Ser Thr Leu
85          90          95
Ile Pro Gly Val Lys Ile Ser Ala Gly Ile Ile Ser Lys Leu Pro Trp
100         105         110
Thr Met Lys Thr Met Tyr Glu Ile Val Ser Ser Pro Asn Asn Asn Asn
115         120         125
Asn Asn Gly Asp Tyr Tyr Ser Thr Cys Arg Arg Met Val Met Glu Tyr
130         135         140
Pro Ile Gly Gly Leu Leu His Thr Pro Ala Ile Thr Asn Lys Tyr Pro
145         150         155         160
Arg Ser Arg Met Val Thr Cys Thr Lys Gly Lys Asp His Gln Lys Leu
165         170         175
Tyr Asp Ile Ser Arg Gln Met Phe Asp Ile Ile Glu Ala Asn Gly Gln
180         185         190
Leu

```

&lt;210&gt; 82

&lt;211&gt; 615

&lt;212&gt; DNA

&lt;213&gt; SHRIMP

&lt;400&gt; 82

```

atggatcttt ctttcaactct ttcgggtcgtg tcggccatcc tcgccatcac tgctgtgatt 60
gctgtattta ttgtgatttt taggtatcac aacactgtga ccaagaccat cgaaacccac 120
acagacaata tcgagacaaa catggatgaa aacctccgca ttctgtgtac tgctgaggtt 180
ggatcaggct acttcaagat gactgatgtg tcctttgaca gcgacacctt gggcaaaatc 240
aagatccgca atggaaagtc tgatgcacag atgaaggaag aagatgcgga tctgtcatc 300
actcccgtag agggccgagc actcgaagtg actgtggggc agaattctac ctttgaggga 360
acattcaagg tgtggaacaa cacatcaaga aagatcaaca tcaactggtat gcagatgggtg 420
caaagatta acccatcaaa ggcctttgtc ggtagctcca acacctcctc cttcaccccc 480
gtctctattg atgaggatga agttggcacc tttgtgtgtg gtaccacctt tggcgcacca 540
attgcagcta ccgcgggtgg aaatcttttc gacatgtacg tgcacgtcac ctactctggc 600
actgagaccg agtaa 615

```

&lt;210&gt; 83

```
<211> 204
<212> PRT
<213> SHRIMP
```

<400> 83															
Met	Asp	Leu	Ser	Phe	Thr	Leu	Ser	Val	Val	Ser	Ala	Ile	Leu	Ala	Ile
1				5					10					15	
Thr	Ala	Val	Ile	Ala	Val	Phe	Ile	Val	Ile	Phe	Arg	Tyr	His	Asn	Thr
			20					25					30		
Val	Thr	Lys	Thr	Ile	Glu	Thr	His	Thr	Asp	Asn	Ile	Glu	Thr	Asn	Met
		35					40					45			
Asp	Glu	Asn	Leu	Arg	Ile	Pro	Val	Thr	Ala	Glu	Val	Gly	Ser	Gly	Tyr
	50					55					60				
Phe	Lys	Met	Thr	Asp	Val	Ser	Phe	Asp	Ser	Asp	Thr	Leu	Gly	Lys	Ile
65				70						75					80
Lys	Ile	Arg	Asn	Gly	Lys	Ser	Asp	Ala	Gln	Met	Lys	Glu	Glu	Asp	Ala
			85						90					95	
Asp	Leu	Val	Ile	Thr	Pro	Val	Glu	Gly	Arg	Ala	Leu	Glu	Val	Thr	Val
			100					105					110		
Gly	Gln	Asn	Leu	Thr	Phe	Glu	Gly	Thr	Phe	Lys	Val	Trp	Asn	Asn	Thr
		115					120					125			
Ser	Arg	Lys	Ile	Asn	Ile	Thr	Gly	Met	Gln	Met	Val	Pro	Lys	Ile	Asn
		130				135					140				
Pro	Ser	Lys	Ala	Phe	Val	Gly	Ser	Ser	Asn	Thr	Ser	Ser	Phe	Thr	Pro
145				150						155					160
Val	Ser	Ile	Asp	Glu	Asp	Glu	Val	Gly	Thr	Phe	Val	Cys	Gly	Thr	Thr
			165						170					175	
Phe	Gly	Ala	Pro	Ile	Ala	Ala	Thr	Ala	Gly	Gly	Asn	Leu	Phe	Asp	Met
			180					185					190		
Tyr	Val	His	Val	Thr	Tyr	Ser	Gly	Thr	Glu	Thr	Glu				
		195					200								

<210>	84
<211>	888
<212>	DNA
<213>	SHRIMP

<400>	84					
atgaagaatt	cgcggcagag	gtcaggcggtg	tggagggggta	actcgtgttt	gtacaaatca	60
ttttattttt	ctggtgccat	aattgaatgt	aaaaaaataa	ggataataat	gatgttcctt	120
ttattattctc	tcatactttt	tgtgtgtttt	tgtgggtgtg	tggttgggtg	gattttcatg	180
tcacgaccaa	ataaaactac	tacaacatcc	ataaaaaaaa	caagagaaga	taaagagaag	240
gaaaaagaag	atgacaccga	aggagctgta	ttagggcgaa	gggaacctga	aaataggccg	300
atcggaagag	acgaggaagg	tgctgtagaa	gacggaaaag	aagaggagga	agttttgaa	360
tttgaacaac	cgaagtgtaaa	tactgggtcc	aatacaggag	gtggaagga	aggaactgtg	420
cctggagaag	gtttgtttacc	tccacccctt	cctactccta	ctcctacttc	tccacctact	480
ctctactcta	ctctctcacc	tcccccgaca	cgaaccccat	ctcctttctt	atctcttggg	540
gaagatgatg	atgatgatat	agacatagac	tttgatgata	atgatataga	cgaattttta	600
gatagtggag	aagaaatgga	agaagacgaa	gaagaggaag	atttggaac	actcctttca	660
agactagaaa	caggcatgag	cggcgaagaa	gtagattttg	atgcatcatc	tgcatatatt	720
caaccaggatc	ctgtagttgt	caaaaacata	gaaaggtcag	attatactct	ggacccaatg	780
gagtcgtgtg	agtttttgaa	catagctgag	ggagatatta	gattcttcgt	agatcgaggg	840
ataaccaaca	agattaaagc	catgacggaa	gactctgaag	aactgtaa		888

```
<210> 85
<211> 295
<212> PRT
<213> SHRIMP
```

<400> 85

WO 01/38351

170

PCT/US00/28888

Met	Lys	Asn	Ser	Arg	Gln	Arg	Ser	Gly	Val	Trp	Arg	Gly	Asn	Ser	Cys
1				5					10					15	
Leu	Tyr	Lys	Ser	Phe	Tyr	Phe	Ser	Gly	Ala	Ile	Ile	Glu	Cys	Lys	Lys
			20					25					30		
Ile	Arg	Ile	Ile	Met	Met	Phe	Leu	Leu	Leu	Ser	Leu	Ile	Leu	Phe	Val
		35					40					45			
Cys	Phe	Val	Gly	Val	Val	Val	Gly	Val	Ile	Phe	Met	Ser	Arg	Pro	Asn
	50					55					60				
Lys	Thr	Thr	Thr	Thr	Ser	Asn	Lys	Lys	Thr	Lys	Lys	Asp	Lys	Glu	Lys
65					70					75				80	
Glu	Lys	Glu	Asp	Asp	Thr	Glu	Gly	Ala	Val	Leu	Gly	Arg	Arg	Glu	Pro
			85					90						95	
Glu	Asn	Arg	Pro	Ile	Gly	Arg	Asp	Glu	Glu	Gly	Ala	Val	Glu	Asp	Gly
			100					105					110		
Lys	Glu	Glu	Glu	Glu	Val	Phe	Glu	Phe	Glu	Gln	Pro	Ser	Val	Asn	Thr
		115					120					125			
Gly	Ser	Asn	Thr	Gly	Gly	Gly	Gly	Thr	Gly	Thr	Val	Pro	Gly	Glu	Gly
	130				135						140				
Leu	Leu	Pro	Pro	Pro	Pro	Thr	Pro	Thr	Pro	Thr	Pro	Pro	Pro	Pro	Thr
145					150					155					160
Pro	Thr	Pro	Thr	Pro	Pro	Pro	Pro	Pro	Thr	Arg	Thr	Pro	Ser	Pro	Ser
			165						170					175	
Ser	Ser	Leu	Gly	Glu	Asp	Asp	Asp	Asp	Asp	Ile	Asp	Ile	Asp	Phe	Asp
		180					185					190			
Asp	Asn	Asp	Ile	Asp	Glu	Phe	Leu	Asp	Ser	Gly	Glu	Glu	Met	Glu	Glu
		195					200					205			
Asp	Glu	Glu	Glu	Glu	Asp	Leu	Asp	Thr	Leu	Leu	Ser	Arg	Leu	Glu	Thr
	210					215					220				
Gly	Met	Ser	Gly	Glu	Glu	Val	Asp	Phe	Asp	Ala	Ser	Ser	Ala	Tyr	Ile
225					230					235					240
Gln	Pro	Asp	Pro	Val	Val	Val	Lys	Asn	Ile	Glu	Arg	Ser	Asp	Tyr	Thr
			245						250					255	
Leu	Asp	Pro	Met	Glu	Ser	Trp	Lys	Val	Leu	Asn	Arg	Ser	Glu	Gly	Asp
			260					265					270		
Ile	Arg	Phe	Phe	Val	Asp	Arg	Gly	Ile	Thr	Asn	Lys	Ile	Lys	Ala	Met
	275						280					285			
Thr	Glu	Asp	Leu	Lys	Glu	Leu									
	290					295									

<210> 86  
 <211> 603  
 <212> DNA  
 <213> SHRIMP

<400> 86  
 atgaccatgt ggaacaaaat tgtaataact actaaacgaa tgaattggcc tatggttggt 60  
 ggtgttttct ttattctagc aataactgcc ttagctgtcg ctctctatat acgacatgct 120  
 tctaaacaag aaaaatactc cacttctcat atcaatgaac aatttactgc caaacagttg 180  
 ccagtcactt acttggtcaaa aacgggtaaa ttaaaggaca tgcacttgac ccattccgat 240  
 tttatggcat atgttgatgt gcacaataga acaaaaaactc tgaagcacc catgtgtact 300  
 gacgaggtg gctgggcccc cttttgcctg ctggcttctg ctgaagctta tcgtcgcatc 360  
 cgttatggga gaggagaatt tggaccagaa aaacattccc tagctgaaac tatccaatcg 420  
 acagtacaag atatgtcgga gccatatata acccacatct tcaagaagaa cacggatggt 480  
 gatggacacg gaatgcaatc tgttttggag aagaatagga ataaaatcag aatgggtgat 540  
 ggaaaaacgt catcagaaac ttataatctt agtgataagt ctatatctat tgttgggtgta 600  
 tag 603

<210> 87  
 <211> 196  
 <212> PRT

WO 01/38351

171

PCT/US00/28888

<213> SHRIMP

<400> 87

Met	Thr	Met	Trp	Asn	Lys	Ile	Val	Ile	Thr	Thr	Lys	Arg	Met	Asn	Trp
1				5					10					15	
Pro	Met	Val	Val	Gly	Val	Phe	Phe	Ile	Leu	Ala	Ile	Thr	Ala	Val	Tyr
			20					25					30		
Ile	Arg	His	Ala	Ser	Lys	Gln	Glu	Lys	Tyr	Ser	Thr	Ser	His	Ile	Asn
		35					40					45			
Glu	Gln	Phe	Thr	Ala	Lys	Gln	Leu	Pro	Val	Thr	Tyr	Leu	Ser	Lys	Thr
		50				55					60				
Gly	Lys	Leu	Lys	Asp	Met	His	Leu	Thr	His	Ser	Asp	Phe	Met	Ala	Tyr
65					70				75					80	
Val	Asp	Val	His	Asn	Arg	Thr	Lys	Thr	Leu	Lys	His	Pro	Met	Cys	Thr
				85					90					95	
Asp	Glu	Ala	Gly	Trp	Ala	His	Phe	Cys	Leu	Leu	Ala	Ser	Ala	Glu	Ala
			100					105					110		
Tyr	Arg	Arg	Ile	Arg	Tyr	Gly	Arg	Gly	Glu	Phe	Gly	Pro	Glu	Lys	His
		115				120						125			
Ser	Leu	Ala	Glu	Thr	Ile	Gln	Ser	Thr	Val	Gln	Asp	Met	Ser	Glu	Pro
		130				135					140				
Tyr	Ile	Thr	His	Ile	Phe	Lys	Lys	Asn	Thr	Asp	Val	Asp	Gly	His	Gly
145					150					155				160	
Met	Gln	Ser	Val	Leu	Glu	Lys	Asn	Arg	Asn	Lys	Ile	Arg	Met	Gly	Asp
				165					170					175	
Gly	Lys	Thr	Ser	Ser	Glu	Thr	Tyr	Asn	Leu	Ser	Asp	Lys	Ser	Ile	Ser
			180					185					190		
Ile	Val	Gly	Val												
			195												

<210> 88

<211> 861

<212> DNA

<213> SHRIMP

<400> 88

atgagctcgg	gtagtatcaa	caaccatccc	tcctcaaaca	tggaacaccaa	caaaatggag	60
gaaggggaag	aacaggattt	tgatgtatta	gaactggact	actccaaaat	catccatgat	120
attactgccca	tggtatctgt	tgcagctcct	cctcccaatt	ccatactgga	cgcatctgat	180
ggactaatag	caactgcgctc	tgcgacagca	ccagcagccg	aaacaggaaa	ctctaacagg	240
atgaggctag	ataaagacgt	gtgccaaacta	atcgaaagg	acatagaatt	agtgaagagt	300
gtactatttg	aagttgactc	cattattcgc	caactgttat	attttggaga	atctgcatca	360
gagaaaaaca	taaaaaccaa	ctctactgaa	aaggagccag	tttacttccc	caaagaaccg	420
aaaggggagg	cagtcaaact	ggctaagaat	accccgattc	tagatacgat	aacaaaacta	480
gattggatgg	cgaacatctg	ccaaagcaac	aagatcgggg	ttgaaaattt	agcctctgcg	540
ttacaaagtg	ggcaattaat	atggacaact	ttcccagctg	ctgtatatgc	ttcgttggac	600
agcttttatc	acattgctat	aatgtggaaa	cttttaggtt	cctttataaa	catcgaagct	660
ttatcaaaag	gatcaaagga	caatcttttg	cctcgtgatg	atatacaggt	tgtccatgct	720
aaacaggaga	ttgctgcaat	gcttcaatct	agacaaaata	ttctcggaag	aggaccgctg	780
gaatatccgc	ctgtgcctat	cactgcaatt	ctatcacgta	ctataattcc	cctgttgagg	840
aacttttctg	agaaattatg	a				861

<210> 89

<211> 286

<212> PRT

<213> SHRIMP

<400> 89

Met	Ser	Ser	Gly	Ser	Ile	Asn	Asn	His	Pro	Ser	Ser	Asn	Met	Asp	Thr
1				5					10					15	

```

Asn Lys Met Glu Glu Gly Glu Glu Gln Asp Phe Asp Val Leu Glu Leu
      20      25      30
Asp Tyr Ser Lys Ile Ile His Asp Ile Thr Ala Met Leu Ser Val Ala
      35      40      45
Ala Pro Pro Pro Asn Ser Ile Leu Asp Ala Ser Asp Gly Leu Ile Ala
      50      55      60
Thr Ala Ser Ala Thr Ala Pro Ala Ala Glu Thr Gly Asn Ser Asn Arg
      65      70      75      80
Met Arg Leu Asp Lys Asp Val Cys Gln Leu Ile Glu Arg Asp Ile Glu
      85      90      95
Leu Val Lys Ser Asp Thr Ile Glu Val Asp Ser Ile Ile Arg Gln Leu
      100     105     110
Leu Tyr Phe Gly Glu Ser Ala Ser Glu Lys Asn Ile Lys Thr Asn Ser
      115     120     125
Thr Glu Lys Glu Pro Val Tyr Phe Pro Lys Glu Pro Lys Gly Glu Ala
      130     135     140
Val Lys Leu Ala Lys Asn Thr Pro Val Leu Asp Thr Ile Thr Lys Leu
      145     150     155     160
Asp Trp Met Ala Asn Ile Cys Gln Ser Asn Lys Ile Gly Val Glu Asn
      165     170     175
Leu Ala Ser Ala Leu Gln Ser Gly Gln Leu Ile Trp Thr Thr Phe Pro
      180     185     190
Ala Ala Val Tyr Ala Ser Leu Asp Ser Phe Tyr His Ile Ala Ile Met
      195     200     205
Trp Lys Leu Leu Gly Ser Phe Ile Asn Ile Glu Ala Leu Ser Lys Gly
      210     215     220
Ser Lys Asp Asn Leu Leu Pro Arg Asp Asp Ile Gln Val Val His Ala
      225     230     235     240
Lys Gln Glu Ile Ala Ala Met Leu Gln Ser Arg Gln Asn Ile Leu Gly
      245     250     255
Arg Gly Pro Ser Glu Tyr Pro Pro Val Pro Ile Thr Ala Ile Leu Ser
      260     265     270
Arg Thr Ile Ile Pro Leu Leu Arg Asn Phe Ser Glu Lys Leu
      275     280     285

```

<210> 90  
 <211> 696  
 <212> DNA  
 <213> SHRIMP

```

<400> 90
atggaccctg gagcatcggc tgettccagg agggcctttat ggtctagcac tgtaacaaac 60
acaaggcatt atcagcagca attgaaccgt gcactgaata agattgaaga agaagatgat 120
gtcgaggagg aacatgggca ggtaacaaca acaaacaaaag aaatggcttc tacttctaca 180
tcttcttctc cactctctc ttcttctccc acgtcttcgg ccatcccttc aagcgatgag 240
gaagaagaag aagaagaaga atatgactct gaatcagaca ctaacgtcga ttctcttctt 300
ggagaggagg aggaagaaga ttcagatata gaatccacat ctgctgatgc aaactttcta 360
cgatcttctt cacggaattc cacgaccaga aacaggctaa taaaaaagta cgttgataga 420
tttatcaagt acgaaaagga tattttactt gctgacagga ataaaaggaa gaagaggcac 480
cgtaatcggc aaccccaaata acataagcta aacaacaaaa gactcaaaaa accaacagac 540
aagaaacaga aaaccaataa gaagaaaact tggagaagac tccctaaatt tattaaaaag 600
atgtctcctg caagttagact gaaatttttc tctgcttgta taatatcagg gataaaaata 660
acttctataa tagtactgtc gataatggcc ctttag 696

```

<210> 91  
 <211> 231  
 <212> PRT  
 <213> SHRIMP

<400> 91



WO 01/38351

PCT/US00/28888

173

Met Asp Pro Gly Ala Ser Ala Ala Ser Arg Arg Ala Leu Trp Ser Ser  
 1 5 10 15  
 Thr Val Thr Asn Thr Arg His Tyr Gln Gln Gln Leu Asn Arg Ala Leu  
 20 25 30  
 Asn Lys Ile Glu Glu Glu Asp Asp Val Glu Glu Glu His Gly Gln Val  
 35 40 45  
 Thr Thr Thr Asn Lys Glu Met Ala Ser Thr Ser Thr Ser Ser Ser Ser  
 50 55 60  
 Ser Ser Ser Ser Ser Pro Thr Ser Ser Ala Ile Pro Ser Ser Asp Glu  
 65 70 75 80  
 Glu Glu Glu Glu Glu Glu Glu Tyr Asp Ser Glu Ser Asp Thr Asn Val  
 85 90 95  
 Asp Ser Leu Leu Gly Glu Glu Glu Glu Glu Asp Ser Asp Thr Glu Ser  
 100 105 110  
 Thr Ser Ala Asp Ala Asn Phe Leu Arg Ser Ser Ser Arg Asn Ser Thr  
 115 120 125  
 Thr Arg Asn Arg Leu Ile Lys Lys Tyr Val Asp Arg Phe Ile Lys Tyr  
 130 135 140  
 Glu Lys Asp Ile Leu Leu Ala Asp Arg Asn Lys Arg Lys Lys Arg His  
 145 150 155 160  
 Arg Asn Arg Gln Pro Gln Ile His Lys Leu Asn Asn Lys Arg Leu Lys  
 165 170 175  
 Lys Pro Thr Asp Lys Lys Gln Lys Thr Asn Lys Lys Lys Thr Trp Arg  
 180 185 190  
 Arg Leu Pro Lys Phe Ile Lys Lys Met Ser Pro Ala Ser Arg Leu Lys  
 195 200 205  
 Phe Phe Ser Ala Cys Ile Ile Ser Gly Ile Lys Ile Thr Ser Ile Ile  
 210 215 220  
 Val Leu Ser Ile Met Ala Leu  
 225 230

<210> 92  
 <211> 4608  
 <212> DNA  
 <213> SHRIMP

<400> 92  
 atggctcaca aactttttatt tctggaggaa gaagatgcta aagagattgg caccctctcc 60  
 caccagaac cttcattcgc tctttatgag agtgaaacat tccggtcggg gggcttctgc 120  
 aaaaatgtaa ccgatgccta ccccaaattt cttcctcggc cgatggacat taattctgtg 180  
 caggcactag cagtgcgtct agccctgatt caattctaca aaggagagg gtggaagaag 240  
 aacatgtcca ttatagacct tgtgaaagat aagggtcgaaa ggaattttaa agtggacaag 300  
 aaaacgagtg gaggatttat catcggagat ggtactggtg taggaaaaac aagagaattg 360  
 gcagccttcg tcatgagtgt catactacaa gaaaaggcac tcttagatgt acaaaaaacac 420  
 gtagggccat caatttttgg tcaagattcg gacaaggtaa ttacggccat caacagtggc 480  
 gtgtggaaaa ggcacccctt ttttatatgg ttgacgtgct ctaagcctct tttcaacagc 540  
 tgtcaacagg gcatgcgaga agtgggttaca aattctagag gtcttcgtga tcctaaattt 600  
 tcatggagaa aactacaagt gccttgtgct aataaaccac cgtctttcaa gagcgacggg 660  
 aaaagtgggt ccatgacagt agatgtggaa aattctgtat cgtctgctaa agactcggtg 720  
 gatattagat ttttactct cagggatgtt aagggaatttc atagtaagcg ttcttcaaga 780  
 tcaataggag atttcttaac agaaacgccc accattttat ttatgacgta ctctgattta 840  
 cgacaaaac tggaatttgt actgaaattt ataacaggag gaacagacct cgattccaac 900  
 aagggttatgc ccattgacaa ctttgtcaca gcattattat gtgacgagtt ccacaagact 960  
 caaaatatta gtgactcttt tagaaaggaa ctggcaaaga cttggggagga agaagatact 1020  
 agagttttgc gcaacatata aaaaagggca aacccttcag tgtcagacct tattaatagg 1080  
 ttcaagagtg caatgagtga cgacagaac tttaaagtga agcgcatgaa atcatccaat 1140  
 aataagggcc gagttacaat gtccaactat ttgaaactat tatctcaagc tgatgcgttc 1200  
 cgtattttct tagagatctt aaagtatgat acgttcacgg ttatggccag cgctacgcca 1260  
 ttccagagca atgctgattt gcacatgatt gaccacatac taaggaaaag tgcgccagcc 1320  
 tacacgagca tacaagcctt caaagaagtg tccagtgcta ctccagatgc tatggctgag 1380

```

cacagcgagt acgtgacagt gtttctggaa cagggttatta aactttttaag aaatagagga 1440
caattgggtta gtaggagtat tagcatggct ggagtcgact gctctactac taattgtaaa 1500
gcctcccctc ttcaaaaata cgctatagac gaattggctt cgtactgttt aaacgctaga 1560
caagtactca tcgattctga aaaagtagga ggacatgtca gaagggcctt cacaaagatt 1620
atacgagaac accaagaagg aggaatattg gaagaggaaag atggtgaaaa attagtagca 1680
gagattaatt ccccgctccag aaagaggaaa cgtgctgcaa atgatgatga ttatatgaa 1740
gtaatggaaa atatcgacag gcgtttttaag gtggtggttg tgcgcgatag ggatgtcgca 1800
catgatggga aaacaacact aagaagtata gtacaagatg ctataaaaaac ctactcacag 1860
aaaaaggatg cgcttttctaa tggaggagga ggaataataa cgtctccgga agttgacata 1920
agcagcattg atatggttgc tcaagattta tacgatgcta tcaagaaaaa ggaaaaacct 1980
tccaaaggta aaactgattt taacgaagat tatgacgacg gtgctaataga agaagacgga 2040
tggggtgaag tatttgacga cgagtgtttt gaaaagtga ggagacaata cttcattaat 2100
acagcaagta ctctgtggc agcatgtaaa ggtgcactat tgaatatcaa ggcaacttct 2160
gtcacccgatg cagtgaagcg ttttaaggaca acaaacgaat caaaaaagat ggtcatgtca 2220
ctagagcaaa ctggggattc ttttctcaag aatttgacta cagcatctt acaaacgatc 2280
gccaaggatg aaagtgcgcg aaaatacggc atcgtagacg tggggacatt cgattcatcg 2340
cctgttgcta atacaatatt ctcaggctat agacttttgt gccgtgcagt catgatggct 2400
tcagctttca caatctcctt aaagaataaa accaatagaa gaacttcccc agctcacgtt 2460
atgcttgtgc cttcagtcct tgacacggaa ccgctcatgg cgctagcagg caaccctata 2520
gattcaatca cacaatcaat aggggaggat agtaacgctg aaataacaaa caggaaattg 2580
tgttcgcgta taacgaatag gggattgttt ctggttaaaa acaacactaa aactgccaac 2640
acaaacaaat gtattagtgc attcaacaac acgaaagaag tgcacgttat aatgctggga 2700
cctaaaggca ataccggtct ctctttacat gactcttcta ataactccat gtacgctaaa 2760
cgttaccatt gcgtttttaga tgtaccttat aacgccattg ctttcctaca gactatcggt 2820
aggactcaca gaaacgggtca actttcagtc cctcagttcc tcatattctc tacagactcg 2880
cccgcagaga gaagattttt tgatagtttg gataagagga tcaaagattc aaaggcagga 2940
acatatgccg atagatacag taataattcc gatgatattg ctgccgcagt tatgagggaa 3000
caatttatcg atcaagggtt agtcctaaag actatgggga atatcgtaca gattgtcaca 3060
gccagtatga caaagggtaca cttgatggaa catttttcaa agatgatgat gaggacaaat 3120
agaggagggg ttgcattcgt agaagggttg acgttagaaa atggtatctt cactgaagta 3180
attgtcctcg cgatgcacat agcgctagta gttatagggtg cccaaaataa aatcacatca 3240
tcagacgatc ttgggcattgc tctttctttc acgtctgtac ttctcataa tcaaattattg 3300
tccattgtta aatccgcac tcaattcggt ttttctaatt tgtgtctgca tttagtccac 3360
tttaaactctg attgtgacaa tcttctccca cgcgagaagc gagtgcgtga tgcagcatct 3420
gccctgatag acactttgaa tacgaaaaat aacgaagtta cttccaagac aaataaaaata 3480
gaaagtgatg caccttcatt aactgctcta atgctccctt ctggtcctag gaatagaaag 3540
atggatgttt tctctaatat aatggcatat aataacaata acggcatgga ctttgatgaa 3600
gatgttcccg acaatgatga agacgaaggg tgtttacctt tgcaagaaga aaatgctaca 3660
acgttagctc tatcaaactt ccccatgat tatgatagag ccataaaaaga tgcacatcaa 3720
ctagtgcagg tgagaattgt tggtaagga gagaaagagg gagtgattcc tatcagtga 3780
tgttttagatg tacctgaatt agatatgaca aatctcatac ctgtagtgac ggcaactaat 3840
gttatacaga gtttggttaa agaaaatcct ggtttgcttt ttactataca taatgctgca 3900
ctagcacatt cccacagaga aggttacgga ggttctcatc ttctgggttt ggctaagaaa 3960
ttgtcccgtg gtttcattaa ttttagacag ttccaaaacc aacttttctc acctaaaaaa 4020
gagtctaaga ttatgtacga tatattcttg agcgtcaaag cgatcatggc cagagatgat 4080
aggatgatg gcctttgtga tatgagaatg aacagcatga tggatgcaag tttcttaaa 4140
gtcaggaaga agccagagtg tgttttcatt acaaaattat tggacaagaa ttttaggaga 4200
catataataa atgacgaaga ggaagaaacg agggagaggt ttggcgggga ggaagaagaa 4260
gaagacgacg acgaagaatt tgaagacgag gaggaggaac aagcagaaaag ggagtggggg 4320
gaagaagaag gagaaagtgc ttacgacatt tctgtcatca atgataaaaa taacacaatc 4380
gggcacgatg tggacataat actgtgcaat agaaagaaac taacactcac taaagaaaat 4440
agtgtttttg ttaattgaaca catagactct tttatggttg gaaatttaat aggtgcggaa 4500
ggttctttga taacaatatg ctttgataac tgcacaggcg agtttgaagg tttacccaaa 4560
ttttgcctgt acgactcctc atctaaagac aaagacacca ttccgtag 4608

```

<210> 93  
 <211> 1523  
 <212> PRT  
 <213> SHRIMP

<400> 93

Met 1	Ala	His	Lys	Leu 5	Leu	Phe	Leu	Glu	Glu 10	Glu	Asp	Ala	Lys	Glu 15	Ile
Gly	Thr	Leu	Ser 20	His	Pro	Glu	Pro	Ser 25	Phe	Ala	Leu	Tyr	Glu 30	Ser	Glu
Thr	Phe	Arg 35	Ser	Val	Gly	Phe	Cys 40	Lys	Asn	Val	Thr	Asp 45	Ala	Tyr	Pro
Lys	Phe 50	Leu	Pro	Arg	Pro	Met 55	Asp	Ile	Asn	Ser	Val 60	Gln	Ala	Val	Arg
Leu 65	Ala	Leu	Ile	Gln	Phe 70	Tyr	Lys	Gly	Arg	Gly 75	Trp	Lys	Lys	Asn	Met 80
Ser	Ile	Ile	Asp	Leu 85	Val	Lys	Asp	Lys	Val 90	Glu	Arg	Asn	Phe	Lys 95	Val
Asp	Lys	Lys	Thr 100	Ser	Gly	Gly	Phe	Ile 105	Ile	Gly	Asp	Gly	Thr 110	Gly	Val
Gly	Lys	Thr 115	Arg	Glu	Leu	Ala	Ala 120	Phe	Val	Met	Ser	Val 125	Ile	Leu	Gln
Glu	Lys 130	Ala	Leu	Leu	Asp	Val 135	Gln	Lys	His	Val	Gly 140	Pro	Ser	Ile	Phe
Gly 145	Gln	Asp	Ser	Asp	Lys 150	Val	Ile	Thr	Ala	Ile 155	Asn	Ser	Gly	Val	Trp 160
Lys	Arg	His	Pro	Phe 165	Phe	Ile	Trp	Leu	Thr 170	Cys	Ser	Lys	Pro	Leu 175	Phe
Asn	Ser	Cys	Gln 180	Gln	Gly	Met	Arg	Glu 185	Val	Val	Thr	Asn	Ser 190	Arg	Gly
Leu	Arg	Asp 195	Pro	Lys	Phe	Ser	Trp 200	Arg	Lys	Leu	Gln	Val 205	Pro	Cys	Ala
Asn	Lys 210	Pro	Thr	Ser	Phe	Lys 215	Ser	Asp	Gly	Lys	Ser 220	Gly	Ser	Met	Thr
Val 225	Asp	Val	Glu	Asn 230	Ser	Val	Ser	Ser	Ala	Lys 235	Asp	Ser	Val	Asp	Ile 240
Arg	Phe	Phe	Thr 245	Leu	Arg	Asp	Val	Lys	Glu 250	Phe	His	Ser	Lys 255	Arg	Ser
Ser	Arg	Ser	Ile 260	Gly	Asp	Phe	Leu	Thr 265	Glu	Thr	Pro	Thr 270	Ile	Leu	Phe
Met	Thr	Tyr 275	Ser	Asp	Leu	Arg	Thr 280	Asn	Leu	Glu	Phe	Val 285	Leu	Lys	Phe
Ile	Thr 290	Gly	Gly	Thr	Asp	Leu 295	Asp	Ser	Asn	Lys	Val 300	Met	Pro	Ile	Asp
Asn 305	Phe	Val	Thr	Ala	Leu 310	Leu	Cys	Asp	Glu	Phe 315	His	Lys	Thr	Gln	Asn 320
Ile	Ser	Asp	Ser	Phe 325	Arg	Lys	Glu	Leu	Ala 330	Lys	Thr	Trp	Glu 335	Glu	Glu
Asp	Thr	Arg	Val 340	Leu	Arg	Asn	Ile	Gln 345	Lys	Arg	Ala	Asn	Pro 350	Ser	Val
Ser	Asp	Leu 355	Ile	Asn	Arg	Phe	Lys 360	Ser	Ala	Met	Ser	Asp 365	Asp	Arg	Asn
Phe	Lys 370	Val	Lys	Arg	Met	Lys 375	Ser	Ser	Asn	Asn	Lys 380	Gly	Arg	Val	Thr
Met 385	Ser	Asn	Tyr	Leu 390	Lys	Leu	Leu	Ser	Gln	Ala 395	Asp	Ala	Phe	Arg	Ile 400
Phe	Leu	Glu	Ile 405	Leu	Lys	Tyr	Asp	Thr	Phe 410	Thr	Val	Met	Ala 415	Ser	Ala
Thr	Pro	Phe	Gln 420	Ser	Asn	Ala	Asp	Leu 425	His	Met	Ile	Asp	His 430	Ile	Leu
Arg	Lys	Ser 435	Ala	Pro	Ala	Tyr	Thr 440	Ser	Ile	Gln	Ala	Phe 445	Lys	Glu	Val
Ser	Ser 450	Ala	Thr	Pro	Asp	Ala 455	Met	Ala	Glu	His	Ser 460	Glu	Tyr	Val	Thr
Val 465	Phe	Leu	Glu	Gln 470	Val	Ile	Lys	Leu	Leu	Arg 475	Asn	Arg	Gly	Gln	Leu 480
Val	Ser	Arg	Ser	Ile	Ser	Met	Ala	Gly	Val	Asp	Cys	Ser	Thr	Thr	Asn

					485					490					495
Cys	Lys	Asp	Leu	Gln	Lys	Tyr	Ala	Ile	Asp	Glu	Leu	Ala	Ser	Tyr	Cys
			500					505					510		
Leu	Asn	Ala	Arg	Gln	Val	Leu	Ile	Asp	Ser	Glu	Lys	Val	Gly	Gly	His
		515					520					525			
Val	Arg	Arg	Ala	Phe	Thr	Lys	Ile	Ile	Arg	Glu	His	Gln	Glu	Gly	Gly
	530					535					540				
Ile	Glu	Glu	Asp	Val	Glu	Lys	Leu	Val	Ala	Glu	Ile	Asn	Ser	Pro	Ser
545					550					555					560
Arg	Lys	Arg	Lys	Arg	Ala	Ala	Asn	Asp	Asp	Asp	Leu	Tyr	Glu	Val	Met
			565						570					575	
Glu	Asn	Ile	Asp	Arg	Arg	Phe	Lys	Val	Val	Val	Val	Arg	Asp	Arg	Asp
			580					585				590			
Val	Ala	His	Asp	Gly	Lys	Thr	Thr	Leu	Arg	Ser	Ile	Val	Gln	Asp	Ala
		595					600					605			
Ile	Lys	Thr	Tyr	Ser	Gln	Lys	Lys	Asp	Ala	Leu	Ser	Asn	Gly	Gly	Gly
	610					615					620				
Gly	Ile	Ile	Thr	Ser	Pro	Glu	Val	Asp	Ile	Ser	Ser	Ile	Asp	Met	Val
625					630					635					640
Ala	Gln	Asp	Leu	Tyr	Asp	Ala	Ile	Lys	Lys	Lys	Glu	Lys	Pro	Ser	Lys
			645						650					655	
Gly	Lys	Thr	Asp	Phe	Asn	Glu	Asp	Tyr	Asp	Asp	Gly	Ala	Asn	Glu	Glu
			660					665					670		
Asp	Gly	Trp	Gly	Glu	Val	Phe	Asp	Asp	Glu	Cys	Phe	Glu	Lys	Leu	Arg
		675					680					685			
Arg	Gln	Tyr	Phe	Ile	Asn	Thr	Ala	Ser	Thr	Ser	Val	Ala	Ala	Cys	Lys
		690				695					700				
Gly	Ala	Leu	Leu	Asn	Ile	Lys	Ala	Thr	Ser	Val	Thr	Asp	Ala	Val	Lys
705					710					715					720
Arg	Leu	Arg	Thr	Thr	Asn	Glu	Ser	Lys	Lys	Met	Val	Met	Ser	Leu	Glu
			725						730					735	
Gln	Thr	Gly	Asp	Ser	Phe	Leu	Lys	Asn	Leu	Thr	Thr	Arg	Ile	Leu	Gln
			740					745					750		
Thr	Ile	Ala	Lys	Asp	Glu	Ser	Asp	Ala	Lys	Tyr	Gly	Ile	Val	Asp	Val
		755					760					765			
Gly	Thr	Phe	Asp	Ser	Ser	Pro	Val	Ala	Asn	Thr	Ile	Phe	Ser	Gly	Tyr
		770				775					780				
Arg	Leu	Leu	Cys	Arg	Ala	Val	Met	Met	Ala	Ser	Ala	Phe	Thr	Ile	Ser
785					790					795					800
Leu	Lys	Asn	Lys	Thr	Asn	Arg	Arg	Thr	Ser	Pro	Ala	His	Val	Met	Leu
			805						810					815	
Val	Pro	Ser	Val	Pro	Asp	Thr	Glu	Pro	Leu	Met	Ala	Gly	Asn	Pro	Ile
			820					825					830		
Asp	Ser	Ile	Thr	Gln	Ser	Ile	Gly	Glu	Asp	Ser	Asn	Ala	Glu	Ile	Thr
		835													

Ser	Asn	Asn	Ser	Ile	Asp	Ile	Ala	Ala	Ala	Val	Met	Arg	Glu	Gln	Phe
Ile	Asp	Gln	Gly	Leu	Val	Leu	Lys	Thr	Met	Gly	Asn	Ile	Val	Gln	Ile
Val	Thr	Ala	Ser	Met	Thr	Lys	Val	His	Leu	Met	Glu	His	Phe	Ser	Lys
Met	Met	Met	Arg	Thr	Asn	Arg	Gly	Gly	Val	Ala	Phe	Val	Glu	Gly	Leu
Thr	Leu	Glu	Asn	Gly	Ile	Phe	Thr	Glu	Val	Ile	Val	Leu	Ala	Met	His
Ile	Ala	Leu	Val	Val	Ile	Gly	Ala	Gln	Asn	Lys	Ile	Thr	Ser	Ser	Asp
Asp	Leu	Gly	His	Ala	Leu	Ser	Phe	Thr	Ser	Val	Leu	Pro	His	Asn	Gln
Ile	Leu	Ser	Ile	Val	Lys	Ser	Ala	Ser	Gln	Phe	Val	Phe	Ser	Asn	Leu
Cys	Leu	His	Leu	Val	His	Phe	Lys	Ser	Asp	Cys	Asp	Asn	Leu	Leu	Pro
Arg	Glu	Lys	Arg	Val	Arg	Asp	Ala	Ala	Ser	Ala	Leu	Ile	Asp	Thr	Leu
Asn	Thr	Lys	Asn	Asn	Glu	Val	Thr	Ser	Lys	Thr	Asn	Lys	Ile	Glu	Ser
Asp	Ala	Pro	Ser	Leu	Thr	Ala	Leu	Met	Leu	Pro	Ser	Gly	Pro	Arg	Asn
Arg	Lys	Met	Asp	Val	Phe	Ser	Asn	Ile	Met	Ala	Tyr	Asn	Asn	Asn	Asn
Gly	Met	Asp	Phe	Asp	Glu	Asp	Val	Pro	Asp	Asn	Asp	Glu	Asp	Glu	Gly
Cys	Leu	Pro	Leu	Gln	Glu	Glu	Asn	Ala	Thr	Thr	Leu	Ala	Leu	Ser	Asn
Phe	Pro	His	Asp	Tyr	Asp	Arg	Ala	Ile	Lys	Asp	Ala	His	Gln	Leu	Val
Thr	Val	Arg	Ile	Val	Gly	Gln	Gly	Glu	Lys	Glu	Gly	Val	Ile	Pro	Ile
Ser	Glu	Cys	Leu	Asp	Val	Pro	Glu	Leu	Asp	Met	Thr	Asn	Leu	Ile	Pro
Val	Val	Thr	Ala	Thr	Asn	Val	Ile	Gln	Ser	Leu	Ala	Lys	Glu	Asn	Pro
Gly	Leu	Leu	Phe	Thr	Ile	His	Asn	Ala	His	Ser	His	Arg	Glu	Gly	
Tyr	Gly	Gly	Ser	His	Leu	Leu	Gly	Leu	Ala	Lys	Lys	Leu	Ser	Arg	Gly
Phe	Ile	Asn	Phe	Arg	Gln	Phe	Gln	Asn	Gln	Leu	Phe	Ser	Pro	Lys	Lys
Glu	Ser	Lys	Ile	Met	Tyr	Asp	Ile	Phe	Leu	Ser	Val	Lys	Ala	Ile	Met
Ala	Arg	Asp	Asp	Arg	Tyr	Asp	Gly	Leu	Cys	Asp	Met	Arg	Met	Asn	Ser
Met	Met	Asp	Ala	Ser	Phe	Leu	Lys	Val	Arg	Lys	Lys	Pro	Glu	Cys	Val
Phe	Ile	Thr	Lys	Leu	Leu	Asp	Lys	Asn	Phe	Arg	Arg	His	Ile	Ile	Asn
Asp	Glu	Glu	Glu	Glu	Thr	Arg	Glu	Arg	Phe	Gly	Gly	Glu	Glu	Glu	Glu
Glu	Asp	Asp	Asp	Glu	Glu	Phe	Glu	Asp	Glu	Glu	Glu	Glu	Gln	Ala	Glu
Arg	Glu	Trp	Gly	Glu	Glu	Glu	Gly	Glu	Ser	Ala	Tyr	Asp	Ile	Ser	Val
Ile	Asn	Asp	Lys	Asn	Asn	Thr	Ile	Gly	His	Asp	Val	Asp	Ile	Ile	Leu
Cys	Asn	Arg	Lys	Lys	Leu	Thr	Leu	Thr	Lys	Glu	Asn	Ser	Val	Phe	Val

1460	1465	1470
Asn Glu His Ile Asp Ser Phe Met Val Gly Asn Leu Ile Gly Ala Glu		
1475	1480	1485
Gly Ser Leu Ile Gln Ile Cys Phe Asp Asn Cys Thr Gly Glu Phe Glu		
1490	1495	1500
Gly Leu Pro Lys Phe Cys Leu Tyr Asp Ser Ser Ser Lys Asp Lys Asp		
1505	1510	1515
Thr Ile Pro		1520

<210> 94  
 <211> 2919  
 <212> DNA  
 <213> SHRIMP

<400> 94

atggatgaac	gacgaagaga	tccgttactt	taccctacca	accgaagtat	gcgcttcacc	60
gcgcaaataa	cgttattcgt	gaccgtgttc	gtggttaggct	gtatttttgc	gctagtctgc	120
gctgcaatgg	cctataatgt	agccaaacca	atgtctgtta	attttcaagc	tatacacgaa	180
ttagggatga	aaagtaaatt	gaaagcagtt	caaggagcta	accctgaaaa	gacactagaa	240
gaatacctag	aagccagagg	taggcatgac	ggggtggaag	atgcaagcaa	ttaccacact	300
caccccgctc	tcttgacat	gatgaatttg	actgttaaag	ggaataaatg	gaacgttccg	360
tcagaaacga	aggaaaggaa	ttctcggttc	gaaagtcatg	athtagctgc	aaatagatcc	420
tcattctctt	tcgtgcctga	acaccatata	gatagattga	gtgaggccac	catcgaaaaa	480
tctaacaatt	atctagatgc	tgtatcgggg	aaaaagttaa	gacaaaggat	ggtcaatttg	540
aaggataata	tagaaaaaga	cgatactgaa	ttgtacgatt	ctttatttgg	tgtacatgat	600
atccatcacc	atagtgcac	agggcgtaat	ggtgacgctc	ctctctctcc	tccatctaca	660
tcagaaggac	atgatgaaga	tgtggatatt	ttggcctaca	atactggcgg	gtactgttcc	720
aaccagttc	ctttgaaaga	gggacaaact	tgcacatctg	tgtgtttata	ctcgagagca	780
gttcgtgtga	tgacaccatt	tgttgctgga	ggaacattta	tcacacataa	atctggagaa	840
gatcctaacc	cttattgttg	gtcaggaaac	gtgcccggtg	accacataga	aacatcccca	900
accactggag	aacgtgtagt	taaagagtgt	tctgtgcaca	catctatagt	cgttctgaca	960
gatgatggg	gctggcaatg	cagacctaaa	tacccaacat	atctcgagg	ttcgggagga	1020
acatcgatga	ctgcttgccg	tttcaatcct	tcaacccaca	aaggaccacc	acctccttca	1080
tcttcgacac	caatctatta	tgatgtatta	aagaacaac	aaatacggaa	tcatacagaa	1140
ttccgcaact	cttcttacat	ctccaaattg	cgctcaatctt	cttccttggc	ggaattcaaa	1200
attaaatgca	acgatcctga	atttttatac	aaaaacccca	tcacctgttt	ctgtaataat	1260
aagaaggatg	tactaaataa	cgaccttctt	tctcaagatg	taacaaagga	tatgaaattc	1320
agaggatgt	atgaatgtat	ggaaaacccc	tgtgttatga	tgcccaacat	agaccttcg	1380
ttgtttacgt	ttgatgtgag	cacaatgaag	tgtgtaccgg	gcgttaataa	ttcccaagac	1440
tcaaacaggc	atgcaataat	aggcgatgac	aggacaccgt	tagtggggac	tgttcttgcc	1500
atgggaattt	tcttggccga	ccaatctaaa	agaggggacc	agattcacca	gcaaagacct	1560
aaatcaagta	tcgacgaaac	tactgccaag	aaaattgcac	tgccccaagc	gcccatactc	1620
acgcctttta	acctggatgc	aacaaatact	tctaagaatg	tgctctttgt	accaatacct	1680
agcactgttt	tacctctctt	agaaaatatt	ccccatgtaa	tcattccgcc	gtcttcattg	1740
ttacatagaa	gttggtttgc	tcccgtcctg	aacaaacctt	cctcggggca	gcataggcca	1800
ttctgtactg	caccctttta	cattgagcca	gcagctaattg	ttttggccgg	aaacatacct	1860
caaaaacctt	atgaacatag	tatgttggcc	acagaatgtt	tgcgaaatag	tcgcatgggt	1920
tcgggtagtg	ttcacggcgg	atctgaactt	ttattctcta	cacttttgag	tcaaaataaa	1980
ccctcctctt	atatcagaac	tccccctgga	ggtactccag	ctccagagta	caattctaca	2040
gggtgaccaa	ggctggaaga	aataagggac	ttttttgaac	gtaattttta	cgatgagagg	2100
agattgtctc	agacagaata	cgtaataaag	aaacacgcga	gagggatgag	gacaagtga	2160
atatacttca	aatcaagttc	ttgggatagt	ttaatgaaga	ggaaagaatt	tctaagacat	2220
atcataaaaa	aatctgagga	tacttttgta	ttgaaagagg	ggttattaat	gcgttcatac	2280
ggaccttatg	ctgctactgt	gctagcgcga	gatatgtttg	athtagatta	cctaaaagga	2340
aaacctgcat	caaaaaacttc	cagtacattg	aaagtgtcca	atcctttgca	gtacgctttc	2400
cccacatcct	actctgttct	tcccgaagaa	ggagctaccg	acgatatttt	ttcagtagac	2460
cacaatagga	tatttgacag	cgaaacaatc	cccgactatt	ttgattgctc	caatgtaact	2520
cccgggaagt	aaaaattgtt	cggaacatca	tcttctctcat	ccgagtacag	agtggtatatt	2580
gacgacgatg	cgtgggggtt	gcaatcgttt	cgtttggtac	ataaccccaa	gagtggcccg	2640

```

gtagtacagt cagaccctcg tctagcggtt gatgcgtcaa acatttcctc aacccccgaa 2700
ggagctacaa taacaccctt ctctctattc aagaagtcac tcgtcgagtg ggggcataag 2760
aaggcggacg tgcaagaaac ttctgtgttc agggacggtg tggatacttc tgaggcttat 2820
agacgattac ttgtagagac tagtatggct gtacgtaact cctgggttctc ttttagcatgg 2880
gaaaataaaa actattatatt tgccaaaaat agcagctaa 2919

```

<210> 95  
 <211> 962  
 <212> PRT  
 <213> SHRIMP

<400> 95

```

Met Asp Glu Arg Arg Arg Asp Pro Leu Leu Tyr Pro Thr Asn Arg Ser
 1          5          10          15
Met Arg Phe Thr Ala Gln Ile Thr Leu Phe Val Thr Val Phe Val Leu
          20          25          30
Gly Cys Ile Phe Ala Leu Val Cys Ala Ala Met Ala Tyr Asn Val Ala
          35          40          45
Lys Pro Met Ser Val Asn Phe Gln Ala Ile His Glu Leu Gly Met Lys
          50          55          60
Ser Lys Leu Lys Ala Val Gln Gly Ala Asn Pro Glu Lys Thr Leu Glu
          65          70          75          80
Glu Tyr Leu Glu Arg Arg His Asp Gly Val Glu Asp Asn Tyr Pro Pro
          85          90          95
His Pro Ala Leu Leu Asp Met Met Asn Leu Thr Val Lys Gly Asn Lys
          100          105          110
Trp Asn Val Pro Ser Glu Thr Lys Glu Arg Asn Ser Arg Phe Glu Ser
          115          120          125
His Asp Leu Ala Ala Asn Arg Ser Ser Ser Leu Leu Val Pro Glu His
          130          135          140
His Ile Asp Arg Leu Ser Glu Ala Thr Ile Glu Lys Ser Asn Lys Tyr
          145          150          155          160
Leu Asp Ala Val Ser Gly Lys Lys Phe Arg Gln Arg Met Val Asn Leu
          165          170          175
Lys Asp Asn Ile Glu Lys Asp Asp Thr Glu Leu Tyr Asp Ser Leu Phe
          180          185          190
Gly Val His Asp Ile His His His Ser Ala Ser Gly Val Ser Gly Asp
          195          200          205
Ala Pro Pro Pro Pro Pro Ser Thr Ser Glu Gly His Asp Glu Asp Val
          210          215          220
Asp Ile Leu Ala Tyr Asn Thr Gly Gly Tyr Cys Ser Asn Pro Val Pro
          225          230          235          240
Leu Lys Glu Gly Gln Thr Cys Thr Ser Val Cys Tyr Thr Ser Arg Ala
          245          250          255
Val Arg Val Met Thr Pro Phe Val Ala Gly Gly Thr Phe Ile Thr His
          260          265          270
Lys Ser Gly Glu Asp Pro Lys Pro Tyr Cys Trp Ser Gly Asn Val Pro
          275          280          285
Gly Asp His Ile Glu Thr Ser Pro Thr Thr Gly Glu Arg Val Val Lys
          290          295          300
Glu Cys Ser Val His Thr Ser Ile Val Val Leu Thr Asp Asp Gly Gly
          305          310          315          320
Trp Gln Cys Arg Pro Lys Tyr Pro Thr Tyr Phe Gly Gly Ser Gly Gly
          325          330          335
Thr Ser Met Thr Ala Cys Ala Phe Asn Pro Ser Thr His Lys Gly Pro
          340          345          350
Pro Pro Pro Ser Ser Ser Thr Pro Ile Tyr Tyr Asp Val Leu Lys Lys
          355          360          365
Gln Gln Ile Arg Asn His Thr Glu Phe Arg Asn Ser Ser Tyr Ile Ser
          370          375          380
Lys Leu Arg Gln Ser Ser Ser Leu Ala Glu Phe Lys Ile Lys Cys Asn

```

385	390										395					400	
Asp	Pro	Glu	Phe	Leu	Tyr	Lys	Asn	Pro	Ile	Thr	Cys	Phe	Cys	Asn	Asn		
				405					410					415			
Lys	Lys	Asp	Val	Leu	Asn	Asn	Asp	Leu	Leu	Ser	Gln	Asp	Val	Thr	Lys		
			420					425					430				
Asp	Met	Lys	Phe	Arg	Gly	Met	Tyr	Glu	Cys	Met	Glu	Asn	Pro	Cys	Val		
		435					440					445					
Met	Met	Pro	Asn	Ile	Asp	Pro	Ser	Phe	Val	Thr	Phe	Asp	Val	Ser	Thr		
	450					455					460						
Met	Lys	Cys	Val	Pro	Gly	Val	Asn	Asn	Pro	Gln	Asp	Ser	Asn	Arg	His		
465					470					475				480			
Ala	Ile	Ile	Gly	Asp	Asp	Arg	Thr	Pro	Leu	Val	Gly	Thr	Val	Pro	Ala		
				485					490					495			
Met	Gly	Ile	Phe	Leu	Ala	Asp	Gln	Ser	Lys	Arg	Gly	Asp	Gln	Ile	His		
			500					505					510				
Gln	Gln	Arg	Pro	Lys	Ser	Ser	Ile	Asp	Glu	Thr	Thr	Ala	Lys	Lys	Ile		
		515					520					525					
Ala	Gln	Ala	Pro	Ile	Ile	Thr	Pro	Leu	Asn	Leu	Asp	Ala	Thr	Asn	Thr		
	530					535					540						
Ser	Lys	Asn	Val	Leu	Phe	Val	Pro	Ile	Pro	Ser	Thr	Val	Leu	Pro	Pro		
545					550					555					560		
Leu	Glu	Asn	Ile	Pro	His	Val	Ile	Ile	Arg	Pro	Ser	Ser	Leu	Leu	His		
				565					570					575			
Arg	Ser	Cys	Leu	Ala	Pro	Val	Leu	Asn	Lys	Pro	Ser	Ser	Gly	Gln	His		
			580					585					590				
Arg	Pro	Phe	Cys	Thr	Ala	Pro	Phe	Tyr	Ile	Glu	Pro	Ala	Ala	Asn	Val		
		595					600					605					
Leu	Ala	Gly	Asn	Ile	Pro	Gln	Lys	Pro	Tyr	Glu	His	Ser	Met	Leu	Ala		
	610					615					620						
Thr	Glu	Cys	Leu	Arg	Asn	Ser	Arg	Met	Val	Ser	Gly	Ser	Val	His	Gly		
625					630					635					640		
Gly	Ser	Glu	Leu	Leu	Phe	Ser	Thr	Leu	Leu	Ser	Gln	Asn	Lys	Pro	Ser		
				645					650					655			
Ser	Tyr	Ile	Arg	Thr	Pro	Pro	Gly	Gly	Thr	Pro	Ala	Pro	Glu	Tyr	Asn		
			660					665					670				
Ser	Thr	Gly	Asp	Gln	Arg	Leu	Glu	Glu	Ile	Arg	Asp	Phe	Phe	Glu	Arg		
		675					680					685					
Asn	Phe	Asn	Asp	Glu	Arg	Arg	Leu	Ser	Gln	Thr	Glu	Tyr	Val	Ile	Lys		
	690					695					700						
Lys	His	Arg	Met	Arg	Thr	Ser	Glu	Ile	Tyr	Leu	Lys	Ser	Ser	Ser	Trp		
705					710					715					720		
Asp	Ser	Leu	Met	Lys	Arg	Lys	Glu	Phe	Leu	Arg	His	Ile	Ile	Lys	Lys		
				725					730					735			
Ser	Glu	Asp	Thr	Phe	Val	Leu	Lys	Glu	Gly	Leu	Leu	Met	Arg	Ser	Tyr		
			740					745					75				



[illegible]

<210>	96
<211>	681
<212>	DNA
<213>	SHRIMP

<400>	96						
atggcctcag	tctttgaaga	ccctgctgat	ctctttgcc	acatggactt	gactggcaaa	60	
gttcctacc	gccccaat	cctgttcttc	gaaggcctac	tccccaatc	tggcaaggag	120	
attatcgaga	accgcctgat	ccataaagg	aagtctggag	catttgaagc	agatacaca	180	
ttggcgtatt	tcttccctc	caacaatgaa	gaaaatacga	agaaactcaa	cattggtttc	240	
cagatcaagt	cgaattgcct	gtctttcttt	attagggatt	ttttgaacga	ctggttggag	300	
gagatcaagg	actgtggacc	atactgtact	ttttcccaat	acatggatgg	ggacaaagaa	360	
actctcgga	acagtgtttt	tggacaagac	tttactattg	tcgccatgga	ctggatcgac	420	
aagggtgtaa	cattctatat	atttgttgat	ggatctgatt	ctatggagaa	tatggcctct	480	
ttgtggatgt	gtgacaaact	gaagagaatg	aatgccaatg	tggtaaaagt	gtttgttgat	540	
aatgcactca	agccaaaatt	ttctgtatgt	aaaacatgta	ggtgggagtt	cccaggcca	600	
gtgctgtacg	ttattgaggg	ccacggaatg	ggacattcag	atttgtcatg	tgatgagatt	660	
tctgagtttt	tggtagaata	a				681	

<210>	97
<211>	224
<212>	PRT
<213>	SHRIMP

<400> 97																
Met	Ala	Ser	Val	Phe	Glu	Asp	Pro	Ala	Asp	Leu	Phe	Ala	Asn	Met	Asp	
1				5					10					15		
Leu	Thr	Gly	Lys	Val	Pro	Trp	Asn	Ile	Leu	Phe	Phe	Glu	Gly	Leu	Leu	
			20					25					30			
Pro	Asn	Ser	Gly	Lys	Glu	Ile	Met	Glu	Asn	Arg	Leu	Ile	His	Lys	Gly	
		35					40					45				
Lys	Cys	Gly	Ala	Phe	Glu	Ala	Asp	Thr	Gln	Leu	Ala	Tyr	Phe	Phe	Pro	
	50					55					60					
Ser	Asn	Asn	Glu	Glu	Asn	Thr	Lys	Lys	Leu	Asn	Ile	Gly	Phe	Gln	Ile	
65					70					75					80	
Lys	Ser	Asn	Cys	Leu	Ser	Phe	Phe	Ile	Arg	Asp	Phe	Leu	Asn	Asp	Trp	
				85					90					95		
Leu	Glu	Glu	Ile	Lys	Asp	Cys	Gly	Pro	Tyr	Cys	Thr	Phe	Ser	Gln	Tyr	
			100					105					110			
Met	Asp	Gly	Asp	Lys	Glu	Ile	Phe	Gly	Asn	Ser	Val	Phe	Gly	Gln	Asp	
		115					120					125				
Phe	Thr	Ile	Val	Ala	Met	Asp	Trp	Ile	Asp	Lys	Gly	Val	Thr	Phe	Tyr	
	130					135					140					
Ile	Phe	Val	Asp	Gly	Ser	Asp	Ser	Met	Glu	Asn	Met	Ala	Ser	Leu	Trp	
145					150					155					160	
Met	Cys	Asp	Lys	Leu	Lys	Arg	Met	Asn	Ala	Asn	Val	Val	Lys	Val	Phe	
				165					170					175		

<400>	99														
Met	Ile	Ala	Ile	Ala	Asn	His	Lys	His	Asp	Val	Ser	Asp	Ala	Leu	Val
1				5					10					15	
Gly	Ala	His	Gly	Ala	Lys	Ile	Asn	Met	Leu	Tyr	Gly	Lys	Ser	Ser	Thr
			20					25					30		
Leu	Ser	Val	Thr	Glu	Ala	Ala	Leu	Leu	Met	Phe	Asn	Asp	Thr	Ala	Leu
		35					40					45			
Thr	Gln	Phe	Ala	Gln	Arg	Gly	Tyr	Glu	Pro	Ser	Ile	Pro	Thr	Ile	Leu
.						55					60				
Lys	Ala	Ala	Leu	Asp	Phe	Ser	Leu	Gln	Glu	Glu	Glu	Pro	Leu	Val	Ala
65					70					75					80
Ala	Thr	Gly	Leu	Asp	Val	Asn	Lys	Ala	Pro	Arg	Ser	Trp	Pro	Ile	Leu
				85					90					95	
Asn	Cys	Arg	Leu	Gly	Tyr	Ile	Ala	Ser	Ser	Asn	Tyr	Pro	Trp	Ala	Glu
			100					105					110		
His	Ile	Ile	Ser	Gly	Asp	Lys	Glu	Glu	Ile	Lys	Arg	Ala	Leu	Glu	Glu
		115					120					125			
His	Glu	Lys	Asn	Ala	Asn	Val	Arg	Phe	Asp	Ser	Asp	Asn	Cys	Pro	Val
		130				135					140				
Cys	Leu	Glu	Asp	Phe	Ser	Ser	Thr	Asn	Ile	Ile	Arg	Thr	Thr	Arg	Cys
145					150					155					160
Gly	His	Cys	Ile	Asp	Glu	Lys	Cys	Trp	Asp	Arg	Leu	Val	Leu	Ser	Thr
				165					170					175	
Gln	Arg	Gly	Glu	Ile	Thr	Arg	Cys	Pro	Val	Cys	Arg	Glu	Arg	Thr	Ser
			180					185					190		
Leu	Arg	Pro	Asp	Ala	Asp	Gln	Val	Lys	Glu	Met	Leu	Val	Glu	Pro	Ile
		195					200					205			
Val	Ser	Cys	Lys	Arg	Met	Ala	Val	Pro	Asp	Glu	Gln	Val	Ser	Cys	Lys

<400>	101															
Met	Ala	Glu	Ala	Ala	Pro	Arg	Tyr	Arg	Gln	Val	Leu	Glu	Glu	Val	Leu	
1				5					10					15		
Glu	Asn	Ile	Glu	Pro	Tyr	Met	Ser	Phe	Leu	Asp	Val	Phe	Thr	Glu	Arg	
			20					25					30			
Glu	Leu	Ala	Leu	Leu	Asn	Asp	Ile	Ile	Thr	Ser	Arg	Asn	Ser	Pro	Pro	
		35					40					45				
Val	Pro	Ser	Ser	Ser	Phe	Lys	Lys	Leu	Asp	Asn	Lys	Glu	Glu	Phe	Arg	
	50					55					60					
Asp	Ile	Ile	Tyr	Phe	Phe	Ile	Asn	Asn	Asn	Thr	Lys	Ser	Asp	Ser	Ser	
65					70					75					80	
Pro	Ile	Cys	Glu	Gly	Met	Thr	Phe	Ile	Asn	Ala	Leu	Thr	Thr	Val	Cys	

85 90 95  
 Lys Thr Phe Arg Gly Glu Asn Ile His Asp Asp Phe Leu Phe Val Lys  
 100 105 110  
 Tyr Ser Leu Leu Val Ser Met Asp Asn Gly Phe Leu Arg Arg Glu Thr  
 115 120 125  
 His Gly Ile Lys Phe Gly Thr Gly Asp Asp Ser Arg Thr Gly Phe Lys  
 130 135 140  
 Phe Thr Ser Lys Glu Gln Ala Glu Glu Glu Arg Glu Lys Val Met Arg  
 145 150 155 160  
 Arg Ile Lys Lys Leu Asp Gly Val Leu Ala Ser Leu Lys Lys Ser Thr  
 165 170 175  
 Ser Ser Ala Arg Ser Gly Ile Val Phe Tyr Val Glu Lys Cys Ser Ser  
 180 185 190  
 Val Ile Arg Phe Arg Leu Phe Ser Arg Ile Val Asn Ile Thr Ser Asp  
 195 200 205  
 Tyr Val Ala Glu Met Lys Lys Ser Ala Pro Leu Glu Pro Phe Asp Ile  
 210 215 220  
 Ser Phe Gly Tyr Lys Tyr Phe Val Asp Glu Ser Pro Cys Val Thr Lys  
 225 230 235 240  
 Ala Lys Arg Leu Ile Ser Asn Gly Asn Phe Ile Ile Gly Arg Pro Phe  
 245 250 255  
 Ser Cys Leu Glu Thr Ser Pro Ser Ser Val Ser Thr Asp Phe Arg Glu  
 260 265 270  
 Glu Met Asn Met Asp Ala Arg Ser Ile Ala Arg Leu Asn Trp Thr Asn  
 275 280 285  
 Glu Glu Arg Ala Ser Ala Tyr Arg Ser Val Ile Ile Lys Ser Phe Leu  
 290 295 300  
 Ser Ser Ile Glu Glu Glu Met Val Glu Glu Tyr Cys Glu Thr Thr Thr  
 305 310 315 320  
 Lys Thr Val Ala Glu Met Ala Val Glu Phe Val Asp Val Phe Ile Glu  
 325 330 335  
 Lys Ala Glu Thr Ile Gln His Phe Gln Thr Lys Ile Phe Asp Thr Met  
 340 345 350  
 Pro Lys Phe Ser Ala Glu Met Met Asp Asn Ile Leu Lys Asn Val Ala  
 355 360 365  
 Ile Asn Glu Ala Val Gly Ser Gly Leu Cys Gly Ala Ile Leu Leu Trp  
 370 375 380  
 Met Ile Asn Ser Arg Pro Phe Glu Glu Ile Asp Tyr Asn Tyr Phe Lys  
 385 390 395 400  
 Ile Cys Leu Arg Glu Ile Met Val Arg Lys Lys Thr Asp Lys Leu Cys  
 405 410 415  
 Asp Asn Leu Ile Val Lys Arg Ile Val Ser His Lys Asn Val Val Ile  
 420 425 430  
 Thr Asp Phe Val Lys Gly Tyr Val Arg Leu Cys Val Lys Val Ser Cys  
 435 440 445  
 Phe Met Glu Asp Leu Glu Ala Phe Leu Thr Lys Asn Pro Trp Leu Lys  
 450 455 460  
 His Thr Tyr Phe Asp Glu Lys Gly Asn Thr Leu Leu Cys Tyr Cys Ile  
 465 470 475 480  
 Ile Asn Lys Tyr Ser His Thr Ser Lys Leu Val Lys Gln Glu Lys Leu  
 485 490 495  
 Asn Ile Leu Lys Pro Ser Ala Lys Gly Met Ser Pro Leu Met Val Cys  
 500 505 510  
 Ala Ala Ile Ser Ser Pro Phe Thr Thr Arg Val Gly Ile Glu Ile Leu  
 515 520 525  
 Thr Thr Asn Ser Leu Ala Phe Ser Phe Ile Asn Glu Asn Asn Glu Asn  
 530 535 540  
 Val Phe His Ala Ala Ala Val Ala Thr Ser Cys Asn Phe Leu Asp Ala  
 545 550 555 560  
 Lys Lys Tyr Lys Asn Ile Ile Tyr Asp Phe Asp Arg Ser Ile Val Asn  
 565 570 575

Met 1	Ala	Gly	Thr	Asp 5	Ile	Ile	Ser	Ser	Ser 10	Ser	Ser	Gly	Ser	Ser	Ser
Ser	Lys	Lys	Gly 20	Gly	Cys	Ile	Val	Ser 25	Lys	Lys	Gly	Lys	Thr 30	Ile	Lys
Gly	Lys	Asn 35	Ile	Val	Phe	Lys	Thr 40	Ser	Ile	Lys	Thr	Ser 45	Ser	Ser	Ser
Glu	Met 50	Met	Lys	Lys	His	Lys 55	Lys	Arg	Met	Glu	Ile 60	Lys	Asp	Met	Val
Lys 65	Lys	Cys	Ala	Ser	Cys 70	Lys	Lys	Val	Asp	Tyr 75	Ser	Ser	Ser	Thr	Leu 80
Glu	Asn	Asp	Ala	Leu 85	Arg	Ala	Ser	Ile	Glu 90	Ser	Thr	Cys	Ser 95	Ala	Leu
Asn	Arg	Phe	Pro 100	Glu	Ile	Lys	Tyr	Gly 105	Glu	Gly	Glu	Ile	Gly 110	Asp	Val
Leu	Ser	Ala 115	Ile	Arg	Leu	Met	Ala 120	Gly	Cys	Leu	Leu	Ala 125	Lys	Asn	Glu
Lys	Ser 130	Phe	Tyr	Lys	Phe	Phe 135	Leu	Arg	Gly	Phe	Gln 140	Phe	Asp	Lys	Asn
Gly 145	Phe	Met	Met	Leu	Ser 150	Glu	Gly	Met	Lys	Arg 155	Ile	Glu	Lys	Met	His 160
Thr	Lys	Ile	Ala	Lys 165	Lys	Val	Phe	Gly	Gly 170	Cys	Lys	Ala	Ala 175	Pro	Leu
Lys	Glu	Asp	Arg 180	Glu	Gly	Lys	Ile	Pro 185	Cys	Gln	Glu	Phe	His 190	Lys	Pro

Ser Ser Tyr Lys Gly Glu Tyr Thr Thr Pro Leu Pro Pro Thr Pro Ala  
195 200 205  
Pro Val Lys Val Leu Pro Pro Leu Leu Pro Tyr Lys Asn Val Lys Asn  
210 215 220  
Lys Pro Val Phe Val Pro Asp Leu Ala Val Gly Glu Ala Lys Lys Pro  
225 230 235 240  
Cys Trp Val His Lys Leu Phe Ser Asp Asp Pro Glu Glu Arg Lys Arg  
245 250 255  
Leu Phe Glu Arg His Gln Ala Gly Arg Arg Asp Ala Leu Met Glu Asp  
260 265 270  
Tyr Gly Val Ile Pro Asn Asn Asp Asn Glu Ala Glu Asp Thr Glu Arg  
275 280 285  
Phe Val Ser Asn Ala Leu Glu Tyr Gln Ala Gln Met Leu Glu Leu Leu  
290 295 300  
Asp Thr Ala Asn Met Pro Pro Pro Ala Ser Thr Pro Val Arg Arg Gly  
305 310 315 320  
Arg Thr Arg Ile Val Arg Asp Tyr Asp Asp Val Pro Ser Pro Tyr Ser  
325 330 335  
Ser Pro Leu His Thr Pro Phe Asp Ala Pro Asn Val Asn Leu Asn Pro  
340 345 350  
Gly Ser Gly Arg Met Val Asp Arg Val Arg Asp Gly Arg Arg Asn Thr  
355 360 365  
Ser Arg Arg Thr Ser Ala Val Met Ala Arg Arg Ile Asn Gln Leu Gln  
370 375 380  
His Gln Phe Leu Tyr Tyr Ser Ser Asp Ser Asp Phe  
385 390 395

<210> 104  
<211> 3796  
<212> DNA  
<213> SHRIMP

<400> 104  
atggcgctcct cgggggggatt ctttacagga atagatgacc ttttcaagac agtgattcaa 60  
caagaaaaaac aagagaaaaa taaaccaact caagcaccag aaacagaacc aaaaccaggg 120  
ccatctcaag ctccagatcc agtcccagac ccagtctcta aaacaccaac caatttctgt 180  
cctcctccac ctaatcctct cctcctcctc cctcctcctc ctcttaagcc ctcaagagaa 240  
gaacggctaa agacgtcaaa aatacgttta aacaaagctc ttagtgatat tgttgaagcc 300  
acaaacgagc gtgttgatgc gttgaaagag aaccaagcat taaatacaga atatgacaag 360  
aaggataatt acttccagggt tttaaagtgc tcgataaacac cttctgtacc aacagctatt 420  
ataggcgac acgtgaaaca ggtggccaaa tacaacgaaa atgagtcgtt aaaatttttc 480  
ctcgatataa aaaataagtg ctcttttagtg gccgtccagt tattctctag gcacgataac 540  
agggaccatg agaaccttat actacaaatt gccgtccagt tattctctag gcacgataac 600  
accaaattgcg tggggggcaga aatatgtgtt aaaggcaacg aaaaaaacaa gtttgtaaac 660  
aaactggtgg taaaaaaact ccccaatgca ccattcctcat cttcaactgt gctagaaatt 720  
agaggcgcta ccagaaattt actggagaat aatttcaaca agggagaaaa taacactgtc 780  
aatgaaaaca aggacattcc tccttcagaa cgagccaacc tggacacgac caaggcagaa 840  
atatcgacacg tcttttccac tctacacaga ctggacacta aaaggaagct tttctttaaa 900  
ggcaacactt tttatcaacg aaaaccaaca ttcgataata aattcagggtg gacagaagtt 960  
ataggggtgga cagaaagtga agcatcaaaa caaaccacta aatcgctaga caagccaacg 1020  
gacgacaatt tattcgtgct accccattct ttcaataatt tggcagacca cttacgtttg 1080  
aaatttataaa acgtcctcta taaaaatagt accgcacatc ccggcaaacg aaattaactac 1140  
aagactcaag agacgctaataaatccccag attgattcgg cgaaagagta caagatgggtc 1200  
tttgagaaaa tcgacaagtgt tttggatgtt cttttggcca tagggaagaa tgacaaatac 1260  
acaaaaagca ctgtcatata atatatagagga aagtttagaa ggtatttaaat attctgctac 1320  
gccttttatg ctctaaataa ggcaaaaacat tctcgcgcag tatccccctc accatttaata 1380  
ttcttttaacc ttttctcctt catgtattgt catggtccgt tttctccattc cgccagtttt 1440  
ttgtccacat tgacgttcgt ctatcaacac atgtttttcc ccatgggcac agccgcccc 1500  
tccgtctcag ccaagcggct catggatata gattccgccc taatgaaagg aggaaagggg 1560  
gtgggtgtga gggatttttg ttcaccttca aaaacaagtc tccatacaag aacattgggtg 1620

```

tctttcttag gttttgctga aatggctatg ggaacaatga cggctctctt atctggtgta 1680
gaagtgcgtg tatctccagc tctccaacaa aggatatacta aatccctaga aagatggtgt 1740
gattcagtc tctttatata tttcaccttt gttttattcc acagattcag tgggtcgaaa 1800
aaagtatcac tcgaatcggc gcttcgcctc atcatggggc agacgcacgc ccacacaaat 1860
aaggtgaggg ccgccaagag atgccgaata gaagcagcgg aaatggaagg tgtggaagaa 1920
gaagaggcgg gcctgacact ctcttatgcc catctattgg gtcttcctta ctctatacaa 1980
aaagccctcg gattacctgt ccctaagata aaccctctca tgacagcatc tttcttctcaa 2040
tacaatttag gggattttgt aggcgtggaa caacttctaa aggctaagag agagtttcca 2100
gccgaaggag aaaccgcagg atttctcggc atgtttgata atctagttaa agattctatt 2160
gacaaatact acggcgaagg agccttttca gacgtggttg aaaatgtaaa acaaggcatg 2220
gaacaaaaca caccgtatga cacatcttca gcgttgatga cacctatccc taaagcattc 2280
tacgaagaag aaaaggatgt tccacagcag gaagaaaatt ctacacaaca aagatataga 2340
ttgaatagag acgtggagga atatttaatg gcttctccta tgaagatggt gtttgtgtct 2400
atactcgata aaactaacca aaaagaacgt ttcattgtctg ttggggatat tgcccttctg 2460
gccgtgtggt gcaaaaaggaa cgtactgaaa aaggattgga acgaatacgc tatcgctaaa 2520
ggcaactacg aatggccttg tgctaaaatg tgcaaccatt tacttttagc tgatttagtg 2580
aattttggaa tattaggtga cttgaaaata accaataaac ttgacacaaa taccgacacc 2640
tttcacagag acagtgatag attaccctca gttgcagatc agaaaaaatt tataaaaaac 2700
acatccctat ctgatcgaaa acaattggcc cttgttccact cgtgcgttaa cgtgagcacc 2760
cgaaccacg taggaagagt gactgcaaca tcatgggctg tcgatgcgct tcgtacctat 2820
acaagagggt ataaagacat gtttgccgcc ctatcttcat cgctggatat gtaccatctt 2880
gggcacacga attcagctaa tttgttcca tattttagta gaaattacct atgtaacgaa 2940
caagagaatg gattgtgggg gtatactcgc agaacctctg aaaaattggc caaagaagaa 3000
ttgggaagag gacgttttag gggcctgaat aaggtagggg tggctaaaac agaactggct 3060
gctgcagcca ttgcaatttc ttctgcctta gatatggggg aagtagaagc tgtaatggac 3120
gactcttcta aagtgaagaa aatagcctcc acctgcttaa atgttaatgc agccaaggctc 3180
tcggccgcca gagaaaaggc gagagaagct agtattaaac gtcttcttct gccactaat 3240
gcaccagcag ctggttccatc cagaaacagt aacagggttc tcctcaaaga tttgtggggg 3300
ttcttttctg acccagacaa gcgccagaag cttataaagg gtgaagcagt ttctgtacta 3360
tgtcccaata caggatttct tcatgctgct gttcctgatt ttgttattga gtattccttc 3420
gaaagtgaag cctctatagt gagattacgt ttgagactga ttaaacctga aaaacaagac 3480
gaaatgggat gcccttcaac agctcccgaag cctaataaga agaggaaatt agtaaggaa 3540
aatcaagacg ctgtactgac gttggatgat gaagataaca tcgttaaata caacaaatat 3600
gatatggttg aagacgagga agcgcgtgaa agattacgcc accaggacaa acaatcggtt 3660
attgcagccc gtatcagtaa agtgtgtgag cggaaaaaatc caaagaaaaa acgtcggtta 3720
gaagaccctg aattgcaaag tgtggatgaa caattgatac gggaactggc tgccattgcc 3780
tactgacgag atctaa
3796

```

<210> 105  
 <211> 1255  
 <212> PRT  
 <213> SHRIMP

```

<400> 105
Met Ala Ser Ser Gly Gly Phe Phe Thr Gly Ile Asp Asp Leu Phe Lys
1      5      10      15
Thr Val Ile Gln Glu Lys Gln Glu Lys Asn Lys Pro Thr Gln Ala
20     25     30
Pro Glu Thr Glu Pro Lys Pro Gly Pro Ser Gln Ala Pro Asp Pro Val
35     40     45
Pro Asp Pro Val Pro Lys Thr Pro Thr Asn Phe Cys Pro Pro Pro Pro
50     55     60
Asn Pro Leu Pro Pro Pro Pro Pro Pro Pro Lys Pro Ser Arg Glu
65     70     75     80
Glu Arg Leu Lys Thr Ser Lys Ile Arg Leu Asn Lys Ala Leu Ser Asp
85     90     95
Ile Val Glu Ala Thr Asn Glu Arg Val Asp Ala Leu Lys Glu Asn Gln
100    105    110
Ala Leu Asn Thr Glu Tyr Asp Lys Lys Asp Asn Tyr Phe Gln Val Leu
115    120    125
Lys Cys Ser Ile Thr Pro Ser Val Pro Thr Ala Ile Ile Gly Ala His

```

130		135		140
Val Lys Gln	Val Ala Lys	Ser Ser Glu	Ile Glu Leu	Ala Val Asn Glu
145	150		155	160
Leu Asp Ile	Lys Asn Lys	Cys Ser Leu	Val Tyr Asn	Glu Asn Glu Ser
	165		170	175
Leu Lys Phe	Phe Arg Asp	His Glu Asn	Leu Ile Leu	Gln Ile Ala Val
	180		185	190
Gln Leu Phe	Ser Arg His	Asp Asn Thr	Lys Cys Val	Gly Ala Glu Ile
	195		200	205
Cys Val Lys	Gly Asn Glu	Lys Asn Lys	Phe Val Asn	Lys Leu Val Val
	210		215	220
Lys Lys Leu	Pro Asn Ala	Pro Ser Ser	Ser Ser Thr	Val Leu Glu Ile
	225		230	235
Arg Gly Ala	Thr Arg Asn	Leu Leu Glu	Asn Asn Phe	Asn Lys Gly Glu
	245		250	255
Asn Asn Thr	Val Asn Glu	Asn Lys Asp	Ile Pro Pro	Ser Ala Asn Leu
	260		265	270
Asp Thr Thr	Lys Ala Glu	Ile Ser His	Val Phe Ser	Thr Leu His Arg
	275		280	285
Leu Asp Thr	Lys Arg Lys	Leu Phe Phe	Lys Gly Asn	Thr Phe Tyr Gln
	290		295	300
Arg Lys Pro	Thr Phe Asp	Asn Lys Phe	Arg Trp Thr	Glu Val Ile Gly
	305		310	315
Trp Thr Glu	Ser Glu Ala	Ser Lys Gln	Thr Thr Lys	Ser Leu Asp Lys
	325		330	335
Pro Thr Asp	Asp Asn Leu	Phe Val Leu	Pro His Ser	Phe Asn Asn Leu
	340		345	350
Ala Asp His	Leu Arg Leu	Lys Phe Lys	Asn Val Leu	Tyr Lys Asn Ser
	355		360	365
Thr Ala His	Pro Gly Lys	Arg Asn Tyr	Tyr Lys Thr	Gln Glu Thr Leu
	370		375	380
Ile Asn Pro	Gln Ile Asp	Ser Ala Lys	Glu Tyr Lys	Met Val Phe Ala
	385		390	395
Glu Ile Asp	Lys Cys Leu	Asp Val Leu	Leu Ala Ile	Gly Lys Asn Asp
	405		410	415
Lys Tyr Thr	Lys Ser Thr	Val Ile Gln	Tyr Arg Gly	Lys Phe Arg Arg
	420		425	430
Tyr Leu Ile	Phe Cys Tyr	Ala Phe Tyr	Ala Leu Asn	Lys Ala Lys His
	435		440	445
Ser Arg Ala	Val Ser Pro	Leu Pro Phe	Asn Phe Phe	Asn Leu Phe Ser
	450		455	460
Phe Met Tyr	Cys His Gly	Pro Phe Leu	His Ser Ala	Ser Phe Leu Ser
	465		470	475
Thr Leu Thr	Phe Val Tyr	Gln His Met	Phe Phe Pro	Met Gly Thr Ala
	485		490	495
Ala Pro Ser	Val Ser Ala	Lys Arg Leu	Met Asp Ile	Asp Ser Ala Leu
	500		505	510
Met Lys Gly	Gly Lys Gly	Val Gly Val	Arg Asp Phe	Gly Ser Pro Ser
	515		520	525
Lys Thr Ser	Leu His Thr	Arg Thr Leu	Val Ser Phe	Leu Gly Phe Ala
	530		535	540
Glu Met Ala	Met Gly Thr	Met Thr Ala	Leu Leu Ser	Gly Val Glu Val
	545		550	555
Arg Val Ser	Pro Ala Leu	Gln Gln Arg	Ile Ser Lys	Ser Leu Glu Arg
	565		570	575
Trp Cys Asp	Ser Val Ile	Phe Ile Tyr	Phe Thr Phe	Val Leu Phe His
	580		585	590
Arg Phe Ser	Gly Ala Lys	Lys Val Ser	Leu Glu Ser	Ala Leu Arg Leu
	595		600	605
Ile Met Gly	Gln Thr His	Ala His Thr	Asn Lys Val	Arg Ala Ala Lys
	610		615	620



Arg	Cys	Arg	Ile	Glu	Ala	Ala	Glu	Met	Glu	Gly	Val	Glu	Glu	Glu	Glu
625					630					635					640
Ala	Gly	Leu	Thr	Leu	Ser	Tyr	Ala	His	Leu	Leu	Gly	Leu	Pro	Tyr	Ser
				645					650						655
Ile	Gln	Lys	Ala	Leu	Gly	Leu	Pro	Val	Pro	Lys	Ile	Asn	Pro	Leu	Met
			660					665					670		
Thr	Ala	Ser	Ser	Ser	Gln	Tyr	Asn	Leu	Gly	Asp	Phe	Val	Gly	Val	Glu
		675					680					685			
Gln	Leu	Leu	Lys	Ala	Lys	Arg	Glu	Phe	Pro	Ala	Glu	Gly	Glu	Thr	Ala
		690				695				700					
Gly	Phe	Leu	Gly	Met	Phe	Asp	Asn	Leu	Val	Lys	Asp	Ser	Ile	Asp	Lys
705					710					715					720
Tyr	Tyr	Gly	Glu	Gly	Ala	Phe	Ser	Asp	Val	Val	Glu	Asn	Val	Lys	Gln
				725					730					735	
Gly	Met	Glu	Gln	Asn	Thr	Pro	Tyr	Asp	Thr	Ser	Ser	Ala	Leu	Met	Thr
			740					745					750		
Pro	Ile	Pro	Lys	Ala	Phe	Tyr	Glu	Glu	Glu	Lys	Asp	Val	Pro	Gln	Gln
		755					760					765			
Glu	Glu	Asn	Ser	Thr	Gln	Gln	Arg	Tyr	Arg	Leu	Asn	Arg	Asp	Val	Glu
	770					775					780				
Glu	Tyr	Leu	Met	Asp	Met	Lys	Met	Val	Phe	Val	Ser	Ile	Leu	Asp	Lys
785					790					795					800
Thr	Asn	Gln	Lys	Glu	Arg	Phe	Met	Ser	Val	Gly	Asp	Ile	Ala	Leu	Leu
				805					810					815	
Ala	Val	Trp	Cys	Lys	Arg	Asn	Val	Leu	Lys	Lys	Asp	Trp	Asn	Glu	Tyr
			820					825					830		
Ala	Ile	Ala	Lys	Gly	Asn	Tyr	Glu	Trp	Leu	Gly	Ala	Lys	Met	Cys	Asn
		835					840					845			
His	Leu	Leu	Leu	Ala	Asp	Leu	Val	Asn	Phe	Gly	Ile	Leu	Gly	Asp	Leu
	850					855					860				
Lys	Ile	Thr	Asn	Lys	Leu	Asp	Thr	Asn	Thr	Asp	Thr	Phe	His	Arg	Asp
865					870					875					880
Ser	Asp	Arg	Leu	Pro	Ser	Val	Ala	Asp	Gln	Lys	Lys	Phe	Ile	Lys	Asn
				885				890						895	
Thr	Ser	Leu	Ser	Asp	Arg	Lys	Gln	Leu	Ala	Leu	Val	His	Ser	Cys	Val
			900					905					910		
Asn	Val	Ser	Thr	Arg	Thr	His	Val	Gly	Arg	Val	Thr	Ala	Thr	Ser	Trp
		915					920					925			
Ala	Val	Asp	Ala	Leu	Arg	Thr	Tyr	Thr	Arg	Gly	Asp	Lys	Asp	Met	Phe
	930					935					940				
Ala	Ala	Leu	Ser	Ser	Ser	Leu	Asp	Met	Tyr	His	Leu	Gly	His	Thr	Asn
945					950					955					960
Ser	Ala	Asn	Phe	Val	Pro	Tyr	Phe	Ser	Arg	Asn	Tyr	Leu	Cys	Asn	Glu
			965					970						975	
Gln	Glu	Asn	Gly	Leu	Trp	Gly	Tyr	Thr	Arg	Arg	Thr	Ser	Glu	Lys	Leu
			980					985					990		
Ala	Lys	Glu	Glu	Leu	Gly	Arg	Gly	Arg	Leu	Gly	Gln	Lys	Val	Gly	Val
		995					1000					1005			
Ala	Lys	Thr	Glu	Leu	Ala	Ala	Ala	Ala	Ile	Ala	Ile	Ser	Ser	Ala	Leu
	1010					1015					1020				
Asp	Met	Gly	Glu	Val	Glu	Ala	Val	Met	Asp	Asp	Ser	Ser	Lys	Val	Arg
1025					1030					1035					1040
Lys	Ile	Ala	Ser	Thr	Cys	Leu	Asn	Val	Asn	Ala	Ala	Lys	Val	Ser	Ala
				1045				1050						1055	
Ala	Arg	Glu	Lys	Ala	Arg	Glu	Ala	Ser	Ile	Lys	Arg	Leu	Leu	Leu	Ala
			1060					1065					1070		
Thr	Asn	Ala	Pro	Ala	Ala	Gly	Ser	Ser	Arg	Asn	Ser	Asn	Arg	Phe	Leu
		1075					1080						1085		
Leu	Lys	Asp	Leu	Trp	Gly	Phe	Phe	Ser	Asp	Pro	Asp	Lys	Arg	Gln	Lys
	1090					1095				1100					
Leu	Ile	Lys	Gly	Glu	Ala	Val	Ser	Val	Leu	Cys	Pro	Asn	Thr	Gly	Phe

WO 01/38351

PCT/US00/28888

190

```

1105          1110          1115          1120
Leu His Ala Ala Val Pro Asp Phe Val Ile Glu Tyr Ser Phe Glu Ser
          1125          1130          1135
Glu Thr Ser Ile Val Arg Leu Arg Leu Arg Leu Ile Lys Pro Glu Lys
          1140          1145          1150
Gln Asp Glu Met Val Cys Pro Ser Thr Ala Pro Glu Ala Asn Lys Lys
          1155          1160          1165
Arg Lys Leu Val Arg Asn Asn Gln Asp Ala Val Leu Thr Leu Asp Asp
          1170          1175          1180
Glu Asp Asn Ile Val Lys Tyr Asn Lys Tyr Asp Met Val Glu Asp Glu
1185          1190          1195          1200
Glu Ala Arg Glu Arg Leu Arg His Gln Asp Lys Gln Ser Val Ile Ala
          1205          1210          1215
Ala Arg Ile Ser Lys Val Cys Glu Arg Lys Asn Pro Lys Lys Arg
          1220          1225          1230
Arg Leu Glu Asp Pro Glu Leu Gln Ser Val Asp Glu Gln Leu Ile Arg
          1235          1240          1245
Glu Leu Ala Ala Ile Ala Tyr
          1250          1255

```

<210> 106  
 <211> 1824  
 <212> DNA  
 <213> SHRIMP

```

<400> 106
atgagagacg atactttttaa ccaagaaact gcggtgaaac ttgtacgatg gtatacagag 60
tacgattgtt gttgccattt ggtaaccgc gtggagcgcc ttctaggatc gttcggagga 120
ggcgtggacg ccacgtctgt gcgaagccga ccggtctctt atgaagaaga taagaaggga 180
gataaatgca tacccttttag gataacgtcc cttattgagg gtatactttt ggaaagggct 240
ctaactaaac ccgattttagc tgctgcagct tttgatgtat cagaaaagct ggtgtattgt 300
agttgtaata acactcaagg caattttgat gtctcttcaa tgaccatatg gattgatggc 360
aataatagta aaaagtatga agttacatgc ccgtcatgca ctgtcgagaa aattagtggg 420
ggtgccgaat ctattcacia gaaacccatg tctcttcttg ccttctttta caatctggta 480
gagaaagaag ccttcgccga aagaattgaa ctcaagaaat tgtacctctc cttactaacg 540
ggctcggcag ccggaggagg aggtatgtac aaggacagct cccaacaatc ttccttcaac 600
ggctcttgga cgtcgttggt gttccacaca tctaaaaagg acaagactcg ttttagaggc 660
gaagttttag tcagtaacaa gataaaacac acatcaagat tgcagcctag gtgcgtctgt 720
tccgatctgt tatatgccct atgctccacc actaacaact ctgcatctta cgcgtacaag 780
gcaagaaatt tgtgtgttat tgaaggtggg gaatttttat attttaaata cacaatcttt 840
gaagagaatg gacctttcga ctccaaaaca gaccttcaat cattagttaa taacgagcct 900
gtttctgaga caaattcatc ggcaactggc gcttcttctt cctctttaga agacgacgat 960
gattgttgtg atgatgatga cgatgatgat gatgatgaag acgaaaaaac taagaagaaa 1020
caacccaaga aacaaaccaa gaaacaaaaa acaacaacat caacacttcc acctatcagc 1080
aaaaccaatc acgacaacat gttgatgaat gtacttaaaa aaggagctgt taatggaaaa 1140
cggaaaatga tggattcttt gtcaggaaaa aagggccaac actctaaaaa attgaaaacc 1200
tccgctgctg ctggtggtgg tgcttcatcc gacgttggtg caggagaaaa tgagggaagag 1260
aacaaccctt cttcagttag tctactaac aatagggata gaaaagacta tgtgcttcca 1320
tgccctcaaa tagaagaagt cactattttt tcacaacaca ggatgaacaa taacaagttg 1380
gcagaaagtg tagtcaaaca ttctgttggt ataatggaa attgtttaaa cttggtttgt 1440
actcaacaca gaaaaaagta tatcctgcct cagaaaaata ttcttttttg ccacacttta 1500
gtccagcatg taggatttaa caattttcgc attttgactg gcgtttcttg cttttttgat 1560
agaattgaaa ttgttttttc cgaccaatct gactctgttg tgttgagtaa taatgctgcc 1620
cattcggcta ttctaaggct attgtcatat ataagagaaa actcattgaa gcgaagtgtg 1680
aggactgctt cggtgaaagg aatcgatttt gtcgtgaaat cacaggacac taatataggc 1740
atacctttta gcaacaagga aataagagaa cggcaattat gctcagcttc aacctgaggt 1800
atgctagctg gtttgggaaa atga
          1824

```

<210> 107  
 <211> 603

<212> PRT  
<213> SHRIMP

<400> 107

Met	Arg	Asp	Asp	Thr	Phe	Asn	Gln	Glu	Thr	Ala	Val	Lys	Leu	Val	Arg
1				5					10					15	
Trp	Tyr	Thr	Glu	Tyr	Asp	Cys	Cys	Cys	Pro	Leu	Val	Asn	Arg	Val	Glu
		20						25					30		
Arg	Leu	Leu	Gly	Ser	Phe	Gly	Gly	Gly	Val	Asp	Ala	Thr	Ser	Val	Arg
		35					40					45			
Ser	Arg	Pro	Ala	Leu	Tyr	Glu	Glu	Asp	Lys	Lys	Gly	Asp	Lys	Cys	Ile
	50					55					60				
Pro	Phe	Arg	Ile	Thr	Ser	Leu	Ile	Glu	Gly	Ile	Leu	Leu	Glu	Arg	Ala
65					70					75					80
Leu	Thr	Lys	Pro	Asp	Leu	Ala	Ala	Ala	Ala	Phe	Asp	Val	Ser	Glu	Lys
				85					90					95	
Leu	Val	Tyr	Cys	Ser	Cys	Asn	Asn	Thr	Gln	Gly	Asn	Phe	Asp	Val	Ser
			100					105					110		
Ser	Met	Thr	Ile	Trp	Ile	Asp	Gly	Asn	Asn	Ser	Lys	Lys	Tyr	Glu	Val
		115					120					125			
Thr	Cys	Pro	Ser	Cys	Thr	Val	Glu	Lys	Ile	Ser	Gly	Gly	Ala	Glu	Ser
	130					135					140				
Ile	His	Lys	Lys	Pro	Met	Ser	Leu	Leu	Ala	Phe	Phe	Asn	Asn	Leu	Val
145					150					155					160
Glu	Lys	Glu	Ala	Phe	Ala	Glu	Arg	Ile	Glu	Leu	Lys	Lys	Leu	Tyr	Leu
				165					170					175	
Ser	Leu	Leu	Thr	Gly	Ser	Ala	Ala	Gly	Gly	Gly	Gly	Met	Tyr	Lys	Asp
			180					185					190		
Ser	Ser	Gln	Gln	Ser	Ser	Phe	Asn	Gly	Ser	Trp	Thr	Ser	Leu	Leu	Phe
		195					200					205			
His	Thr	Ser	Lys	Lys	Asp	Lys	Thr	Arg	Leu	Glu	Ala	Glu	Val	Leu	Val
	210					215					220				
Ser	Asn	Lys	Ile	Lys	His	Thr	Ser	Arg	Leu	Gln	Pro	Arg	Cys	Val	Cys
225					230					235					240
Ser	Asp	Leu	Leu	Tyr	Ala	Leu	Cys	Ser	Thr	Thr	Asn	Asn	Ser	Ala	Ser
				245					250					255	
Tyr	Ala	Tyr	Lys	Ala	Arg	Asn	Leu	Cys	Val	Ile	Glu	Gly	Gly	Glu	Phe
		260						265					270		
Leu	Tyr	Phe	Lys	Tyr	Thr	Ile	Phe	Glu	Glu	Asn	Gly	Pro	Phe	Asp	Ser
		275					280					285			
Lys	Thr	Asp	Leu	Gln	Ser	Leu	Val	Asn	Asn	Glu	Pro	Val	Ser	Glu	Thr
	290					295					300				
Asn	Ser	Ser	Ala	Ala	Ser	Ser	Ser	Ser	Leu	Glu	Asp	Asp	Asp	Asp	Cys
305					310					315					320
Cys	Asp	Asp	Asp	Asp	Asp	Asp	Asp	Asp	Asp	Glu	Asp	Glu	Lys	Thr	Lys
				325					330					335	
Lys	Lys	Gln	Pro	Lys	Lys	Gln	Thr	Lys	Lys	Gln	Lys	Thr	Thr	Thr	Ser
		340						345					350		
Thr	Leu	Pro	Pro	Ile	Ser	Lys	Thr	Asn	His	Asp	Asn	Met	Leu	Met	Asn
		355					360					365			
Val	Leu	Lys	Lys	Gly	Ala	Val	Asn	Gly	Lys	Arg	Lys	Met	Met	Asp	Ser
	370					375					380				
Leu	Ser	Gly	Lys	Lys	Gly	Gln	His	Ser	Lys	Lys	Leu	Lys	Thr	Ser	Ala
385					390					395					400
Ala	Ala	Gly	Gly	Gly	Ala	Ser	Ser	Asp	Val	Val	Ala	Gly	Glu	Asn	Glu
				405					410					415	
Glu	Glu	Asn	Asn	Pro	Ser	Ser	Val	Ser	Pro	Thr	Asn	Asn	Arg	Asp	Arg
			420					425					430		
Lys	Asp	Tyr	Val	Leu	Pro	Cys	Pro	Gln	Ile	Glu	Glu	Val	Thr	Ile	Phe
		435					440					445			
Ser	Gln	His	Arg	Met	Asn	Asn	Asn	Lys	Leu	Ala	Glu	Ser	Val	Val	Lys

450	455	460
His Ser Val Val Ile Asn Gly Asn Cys Leu Asn Leu Phe Val Thr Gln		
465	470	475
His Arg Lys Lys Tyr Ile Leu Phe Asn Ile Leu Phe Cys Pro Pro Leu		480
	485	490
Val Gln His Val Gly Phe Asn Lys Phe Arg Ile Leu Thr Gly Val Ser		495
	500	505
Cys Phe Phe Asp Arg Ile Glu Ile Val Phe Ser Asp Gln Ser Asp Ser		510
	515	520
Val Val Leu Ser Asn Asn Ala Ala His Ser Ala Ile Leu Arg Leu Leu		525
	530	535
Ser Tyr Ile Arg Glu Asn Ser Leu Lys Arg Ser Val Arg Thr Ala Ser		540
545	550	555
Val Lys Gly Ile Asp Phe Val Val Lys Ser Gln Asp Thr Asn Ile Gly		560
	565	570
Ile Pro Leu Ser Asn Lys Glu Ile Arg Glu Arg Gln Leu Cys Ser Ala		575
	580	585
Ser Thr Leu Ser Met Leu Ala Gly Leu Gly Lys		590
595	600	

<210> 108  
 <211> 2403  
 <212> DNA  
 <213> SHRIMP

<400> 108  
 atgctcagct tcaaccctga gtatgctagc tggtttgga aaatgatcac cgacccgggg 60  
 gtcattttac ccgtttccaa ggatgttgta ctctttggat cagcaggaca atcagatgtg 120  
 ggaataatga cactggaccc tcacgatttg gatataaaaa tcacatctaa acgcataggt 180  
 gtggaagaaa gactagctca atacaatact ctacctatgg attttacacg ggcaatggaa 240  
 aaggaaactaa ataatagtag aaatatgaaa gagtcaatat tcacgggaat atttttagac 300  
 accggttcgg caatcttcga agacaacatg ttcaacggag gaggttcagc tttgcgctta 360  
 attagatccc ccgctttgaa ttctgctgta ttttcaagca agaactacat catcaaacia 420  
 ttgcccacca taaccaaata tctaaggaga agtcaagcta gagataagca agtggataaa 480  
 acaagagaga agatagtggg ggattctttc agcatactta gtgctatagc tgctcaagta 540  
 atgcaccta caagcggaga gatgacgtac gtccccgatg ggcactgcgt taatgtgtgc 600  
 atgtcagaga ccaatgcttc gtccatctac ttgatcataa acgacccac tggttcggga 660  
 tggaaaatta tgcctaacia tttcaataaa acacttgaaa tgagagacgg tgtaatagat 720  
 agagttagaaa cattagtggg gtttgcggtc aagtgcgtcg catcatcctt gattaaaagg 780  
 ggcatggatt tagtggatat gcaaagaact ataagggtcta tggatttctt cctccagct 840  
 tcttctactt ccaataatac tcctagagta gcgataatga catctggaag tagtactact 900  
 acgggcattg gatccttgct cattcttgca gaagatggat caacacacca ccaatcaag 960  
 ctatcggaat ataggactgg attatccatt actgaaaata atagagaggt gtcttttacg 1020  
 gtagaacctt caatagacgg cgttcaagca gagcatcctc tatcccttc tattcttcag 1080  
 tggttacctc ctctagttaa aaggccagaa gtggtagcag cagcagcagc agcagtagta 1140  
 gaagaagaaa acggggacaa taaaccttct gataaagata acgaagacaa gtacagtgat 1200  
 actgattttt ggtctaattg ccccgtcaca cctctaatta caccacagaa atggagagcg 1260  
 tgcaaaataa acgatcgggc aatgattagt agttggaaaa ataactctagt gaaactccac 1320  
 aaatatgatt ggacaaataa aactacaaag gttgattatt ttgataagat ggctgccttt 1380  
 gttgcctca tgacctttag aaaattccaa gacatactag cggataacta tgttctcct 1440  
 caaacccctt ctcagggaag tgaatacgca gtgacctgt ctaacgtggc tacactcttt 1500  
 actgacgtgt acggttttga atcgaatgga aataagccat tgtttgccct agaacagcta 1560  
 gaaaatgaaa ccgggattga aagcatatac gtcctaaata tcataggaaa tttccctgat 1620  
 ggtaattctg tcagggtcgt cagactggaa aaggaaatga gtttctcttt gaaggcgaag 1680  
 cagtacttta cagaaatggc catacctcct attaataaaa aatgcaaatg gacagataag 1740  
 gccccgtcat ctgtaaagga gtacaagtat ttttgtgatc taacagcacc catttcaaa 1800  
 agacctagaa aagataacia cgacggcgtt gtggagcatt ctgcgttgac ttatacacct 1860  
 aggtgcatat accacactga acgttgttta gtccatcttt actctgagcc agaaaaata 1920  
 acagaacacg tatctttcaa caaggatttg aacatattag aaattggaaa aaatattacc 1980  
 aaccaatacc aaacaaacta caaaagcata ttcgaaattg tggacgttcc cataattgtc 2040



WO 01/38351

194

PCT/US00/28888

370						375						380					
Gly	Asp	Asn	Lys	Pro	Ser	Asp	Lys	Asp	Asn	Glu	Asp	Lys	Tyr	Ser	Asp		
385						390				395					400		
Thr	Asp	Phe	Trp	Ser	Asn	Val	Pro	Val	Thr	Pro	Leu	Ile	Thr	Pro	Lys		
				405					410						415		
Lys	Trp	Arg	Ala	Cys	Lys	Ile	Asn	Asp	Arg	Ala	Met	Ile	Ser	Ser	Trp		
			420					425					430				
Lys	Asn	Asn	Leu	Val	Lys	Leu	His	Lys	Tyr	Asp	Trp	Thr	Asn	Lys	Thr		
	435						440					445					
Thr	Lys	Val	Asp	Tyr	Phe	Asp	Lys	Met	Ala	Ala	Phe	Val	Met	Thr	Phe		
	450					455					460						
Arg	Lys	Phe	Gln	Asp	Ile	Leu	Ala	Asp	Asn	Tyr	Val	Pro	Pro	Gln	Thr		
465					470				475					480			
Pro	Ser	Gln	Gly	Ser	Glu	Tyr	Ala	Val	Thr	Met	Ser	Asn	Val	Ala	Thr		
				485					490					495			
Leu	Phe	Thr	Asp	Val	Tyr	Gly	Phe	Glu	Ser	Asn	Gly	Asn	Lys	Pro	Leu		
			500					505					510				
Phe	Ala	Leu	Glu	Gln	Leu	Glu	Asn	Glu	Thr	Gly	Ile	Glu	Ser	Ile	Tyr		
	515					520					525						
Val	Leu	Asn	Ile	Ile	Gly	Asn	Ser	Pro	Asp	Gly	Asn	Ser	Val	Arg	Val		
	530				535					540							
Val	Arg	Leu	Glu	Lys	Glu	Met	Ser	Phe	Leu	Leu	Lys	Ala	Lys	Gln	Tyr		
545					550				555					560			
Phe	Thr	Glu	Met	Ala	Ile	Pro	Pro	Ile	Asn	Glu	Lys	Cys	Lys	Trp	Thr		
			565					570						575			
Asp	Lys	Ala	Pro	Ser	Ser	Val	Lys	Glu	Tyr	Lys	Tyr	Phe	Cys	Asp	Leu		
			580					585					590				
Thr	Ala	Pro	Ile	Ser	Lys	Arg	Pro	Arg	Lys	Asp	Asn	Asn	Asp	Gly	Gly		
	595						600					605					
Val	Glu	His	Ser	Ala	Leu	Thr	Tyr	Thr	Pro	Arg	Cys	Ile	Tyr	His	Thr		
	610					615					620						
Glu	Arg	Cys	Leu	Val	His	Lys	Glu	Pro	Glu	Lys	Ile	Thr	Glu	His	Val		
625					630				635					640			
Ser	Phe	Asn	Lys	Asp	Leu	Asn	Ile	Ile	Gly	Lys	Asn	Ile	Thr	Asn	Gln		
				645					650					655			
Tyr	Gln	Thr	Asn	Tyr	Lys	Ser	Ile	Phe	Glu	Ile	Val	Asp	Val	Pro	Ile		
			660					665					670				
Ile	Val	Ala	Ser	Met	Ser	Ser	Thr	Lys	Thr	Met	Thr	Val	Asn	Asn	Tyr		
	675						680					685					
Ile	Ile	Ser	Thr	Pro	Ser	Ala	Thr	Thr	Lys	Phe	Val	Gln	Asp	Pro	Pro		
	690					695					700						
Lys	Thr	Gly	Lys	Gln	Leu	Ala	Val	Glu	Glu	Val	Arg	Asn	Phe	Lys			
705					710				715					720			
Leu	Lys	Ser	Val	Leu	Val	Pro	Pro	Pro	Tyr	Phe	Arg	Asp	Asn	Lys	Arg		
				725					730					735			
Asn	Thr	Thr	Leu	Cys	Ser	Gln	Ile	Thr	Glu	Gln	Asn	Cys	Pro	Ser	Ser		
			740					745					750				
Ser	Glu	Gly	Gly	Arg	Phe	Ser	Cys	Pro	Ser	Glu	Ser	Leu	Ile	Leu	Lys		
	755					760						765					
Tyr	Ser	Asn	Leu	Ser	Lys	Lys	Arg	Ala	Leu	Glu	Glu	Ile	Ala	Pro	Glu		
	770				775						780						
Thr	Glu	Thr	Ser	Ile	Leu	Ser	Leu	Ala	Met								
785					790												

<210> 110  
 <211> 1581  
 <212> DNA  
 <213> SHRIMP  
  
 <400> 110



210		215		220
Gly Arg Met Ser Gly Thr Pro Met Asn Pro Lys Asp Met Thr Tyr Phe				
225		230		240
Val Asn Asp Phe Ser Asp Asp Ile Gly Ser Thr Pro Gln Cys Leu Val				
	245		250	255
Ser Asn Ser Asp Ile Leu Asn Lys Arg Glu Glu Trp Ile Ala Val Trp				
	260		265	270
Gly Val Ala Asp Ser Lys Asp Leu Leu Thr Lys His Gln Leu Gly Glu				
	275		280	285
Arg Glu Tyr Gly Ser Glu Gly Arg Arg Arg Asn Pro Gly Val Glu Glu				
	290		295	300
Glu Glu Glu Glu Arg Val Glu Glu Glu Glu Val Glu Val Pro Tyr				
305		310		320
Ile Lys Lys Ser Gly Lys Leu Ile Gly Pro Arg Arg Arg Pro Leu Thr				
	325		330	335
Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Asn Pro Ile Val Arg				
	340		345	350
Glu Val Val Glu Asp Phe Asp Tyr Glu Ser Phe Asn Glu Pro Glu Ile				
	355		360	365
Phe Gly Ser Asn Ser Lys Leu Pro Phe Ile Arg Phe Leu Asp Gln Lys				
	370		375	380
Asn Trp Arg Leu Gly Ile Met Ser Arg Val Ser Ser Ser Ile Ala Asn				
385		390		400
Phe Lys Ile Glu Gln Glu Ser Ser Lys Ala Leu Phe Cys Leu Ala Val				
	405		410	415
Trp Val Gly Asp Glu His Thr Pro Lys Phe Arg Leu Ser Val Trp Lys				
	420		425	430
Asn Trp Lys Pro Phe Thr Ser Ala Pro Ile Ile Val Gln Asn Val Gly				
	435		440	445
Tyr Ser Ser Asp Val Phe Trp His Glu Thr Leu Arg Ser Lys Ile Val				
	450		455	460
Asp Arg Ser Arg Asp Leu Ile Glu Thr Lys Val Thr Lys Lys Ile Gly				
465		470		480
Glu Asp Trp Ala Asn Lys Lys Gln Thr Val Val Ala Met Phe Ile Ser				
	485		490	495
Gly Ile Val Cys Ile Thr Val Thr Val Ile Ser Ile Phe Ser Ile Val				
	500		505	510
Ile Tyr Tyr Lys Ile Lys Met Pro Lys Phe				
	515		520	

<210> 112  
 <211> 627  
 <212> DNA  
 <213> SHRIMP

<400> 112  
 atgtatatct tcgtcgaagg ttcccccttc acagggaaga gttcatggat gtccaagttg 60  
 atagatacag gatcatgtgg aatgtctttc ctcaattttc ttcgtatgaa cacttctgac 120  
 tactacaact ggcctgccga aatcgggaca gaacatctcc agttagggtt cagagaaacc 180  
 agagtgggtg atggaatggt tgaacctgtc cttaaagacct ttgtcgactc gtggaagaaa 240  
 gagcaaggaa aagagagttt gaaggaatat ctggactaca acggccaagt catggagatc 300  
 tacatcgag aatggttgag acaaaggcca ctagccttcc acgtgtttac ctatacagat 360  
 gaagctgtca agagtggatt cttgaacgag gaggatctag atatggatac tgcaaccaag 420  
 tggatggctg aaattattag agagaagagg ggcaatattc aagaaataaa agtgaccctt 480  
 agagtagtct tcaatggcaa tgtttgtagt gcatgtttct ctaacactaa gagaaacttg 540  
 tataactttg gaacaaacta taacaatgtt gtacattgtg atttgtttgtg ccctttttgca 600  
 aggcatagga ttgtacattt cttataa 627

<210> 113  
 <211> 208



WO 01/38351

197

PCT/US00/28888

<212> PRT  
<213> SHRIMP

<400> 113

Met	Tyr	Ile	Phe	Val	Glu	Gly	Ser	Pro	Leu	Thr	Gly	Lys	Ser	Ser	Trp
1				5					10					15	
Met	Ser	Lys	Leu	Ile	Asp	Thr	Gly	Ser	Cys	Gly	Met	Ser	Phe	Leu	Asn
			20					25					30		
Phe	Leu	Arg	Met	Asn	Thr	Ser	Asp	Tyr	Tyr	Asn	Trp	Pro	Ala	Glu	Ile
		35					40					45			
Gly	Thr	Glu	His	Leu	Gln	Leu	Gly	Phe	Arg	Glu	Thr	Arg	Val	Val	Asp
	50				55					60					
Gly	Met	Phe	Glu	Pro	Val	Leu	Lys	Thr	Phe	Val	Asp	Ser	Trp	Lys	Lys
65					70					75				80	
Glu	Gln	Gly	Lys	Glu	Ser	Leu	Lys	Glu	Tyr	Leu	Asp	Tyr	Asn	Gly	Gln
				85					90					95	
Val	Met	Glu	Ile	Tyr	Ile	Ala	Glu	Trp	Leu	Arg	Gln	Arg	Pro	Leu	Ala
			100					105					110		
Phe	His	Val	Phe	Thr	Tyr	Thr	Asp	Glu	Ala	Val	Lys	Ser	Gly	Phe	Leu
	115						120					125			
Asn	Glu	Glu	Asp	Leu	Asp	Met	Asp	Thr	Ala	Thr	Lys	Trp	Met	Ala	Glu
	130					135					140				
Ile	Ile	Arg	Glu	Lys	Arg	Gly	Asn	Ile	Gln	Glu	Ile	Lys	Val	Thr	Pro
145					150					155				160	
Arg	Val	Val	Phe	Asn	Gly	Asn	Val	Cys	Ser	Ala	Cys	Phe	Ser	Asn	Thr
				165				170						175	
Lys	Arg	Asn	Leu	Tyr	Asn	Phe	Gly	Thr	Asn	Tyr	Asn	Asn	Val	Val	His
		180						185					190		
Cys	Asp	Leu	Leu	Cys	Pro	Phe	Ala	Arg	His	Arg	Ile	Val	His	Phe	Leu
		195					200					205			

<210> 114  
<211> 375  
<212> DNA  
<213> SHRIMP

<400> 114

atgtggagat	cttgtattag	caacattaga	gaaatggggg	acaataaaga	ttatgaaaca	60
agactgattc	aaagaattaa	tgatttggaa	tcagaaattg	aaaataaaac	tgaattatgt	120
gaaaaaatca	atgagcagat	gaaaaatata	caactaaaat	atgataaatg	tttcgtagag	180
gaggagacgg	aaaaattccg	caagatggag	gaaagagtta	aatacctcaa	agagcaggga	240
atccctctag	acccagaaga	aagacgtaca	atgttggctg	aaattgacaa	gagtaacaaa	300
gagttagatg	cccttcttga	ggaaaatgaa	cgtataataa	agctcattga	tgaagagttg	360
gaaagtatga	aataa					375

<210> 115  
<211> 124  
<212> PRT  
<213> SHRIMP

<400> 115

Met	Trp	Arg	Ser	Cys	Ile	Ser	Asn	Ile	Arg	Glu	Met	Gly	Asp	Asn	Lys
1				5					10					15	
Asp	Tyr	Glu	Thr	Arg	Leu	Ile	Gln	Arg	Ile	Asn	Asp	Leu	Glu	Ser	Glu
			20					25					30		
Ile	Glu	Asn	Lys	Thr	Glu	Leu	Cys	Glu	Lys	Ile	Asn	Glu	Gln	Met	Lys
		35					40				45				
Asn	Thr	Gln	Leu	Lys	Tyr	Asp	Lys	Cys	Phe	Val	Glu	Glu	Glu	Thr	Glu
	50					55					60				
Lys	Phe	Arg	Lys	Met	Glu	Glu	Arg	Val	Lys	Tyr	Leu	Lys	Glu	Gln	Gly

WO 01/38351

198

PCT/US00/28888

```

65          70          75          80
Ile Pro Leu Asp Pro Glu Glu Arg Arg Thr Met Leu Ala Glu Ile Asp
          85          90          95
Lys Ser Asn Lys Glu Leu Asp Ala Leu Leu Glu Glu Asn Glu Arg Ile
          100          105          110
Ile Lys Leu Ile Asp Glu Glu Leu Glu Ser Met Lys
          115          120

```

<210> 116  
 <211> 252  
 <212> DNA  
 <213> SHRIMP

```

<400> 116
atgcaaaaaa aatatgataa attatttgaa gatgataaaa gggtccgaga aatagaggaa 60
cgaatccttc aacaaaaaga gaagggaac cctctagacc cagaagaaag acttgatttg 120
tcggctgata ttgataggag tatgaaagag attgatgatt gtctcgagga aataaaccat 180
atagaattat ccattgatac attattggat gaatgtgaaa acttgcatta tggctcttcaa 240
acaactaaat aa                                     252

```

<210> 117  
 <211> 83  
 <212> PRT  
 <213> SHRIMP

```

<400> 117
Met Gln Lys Lys Tyr Asp Lys Leu Phe Glu Asp Asp Lys Arg Phe Arg
 1          5          10          15
Glu Ile Glu Glu Arg Ile Leu Gln Gln Lys Glu Lys Gly Asn Pro Leu
          20          25          30
Asp Pro Glu Glu Arg Leu Val Leu Ser Ala Asp Ile Asp Arg Ser Met
          35          40          45
Lys Glu Ile Asp Asp Cys Leu Glu Glu Ile Asn His Ile Glu Leu Ser
          50          55          60
Ile Asp Thr Leu Leu Asp Glu Cys Glu Asn Leu His Tyr Gly Leu Gln
65          70          75          80
Thr Thr Lys

```

<210> 118  
 <211> 2253  
 <212> DNA  
 <213> SHRIMP

```

<400> 118
atggaaaaaa agactgagac ggctgcaaca acagaaaaag acccagaacc gtctgtcagt 60
aaaagggtcca gaaataaaga acccaaaaca acttctactg tttacacttc tgtaaaagtgt 120
tacctttctt ccataatcaa gagtgaaggt agtagaagta atgtcacctc aaccaaagaa 180
aggtttgagg agagggtgtaa atccgtaagc aagatgatgg tcaaagggtc actggttttg 240
aggtttagtag tggacgagtg tctgagacgt tacaaccatc tagaagacga aatcgataaa 300
tggccagata tgacgaagga taacttttac gtccaattgt tgaggaaggg tttagacaag 360
aagaaattga aagaaggatc tacacatcct gttgtagaag atgttttgaa ttcccccatc 420
gtccaagaaa cattcctatc ccagcaagga gaaggaaata atcccataaa gagacatctc 480
atggatttca ataccatcac ctacgccgcc aaacaactaa aaacttgctt cgaaacaaac 540
ctacgcaccc atttcgggac acgacaacag agggccatat ctggatgggt agctgaaaac 600
gggttcgata aaaagtatac gaaactcgta caaactgga taattggatg tacctacaag 660
agtgattggg tggacagtggt tgatttggaa agggtaaaaag aaggaacgaa aaatttcgtg 720
actcttcata ggaaacattt atgtgttatt agtgataaga agaattggtac aatttcctat 780
tcacctgaag agaaatatcc gataccctca atactaaatt attacaagtt tctacaaaca 840

```

gagtatccac aaaacaagaa aatacagaaa atgatagttg tcccaaaaca caaactaaag 900  
 atacactatt gtacgtttga ccaaacgacc attcaaggaa tttgtaaaga tttgggagtg 960  
 tggaaggata tggaagaacg acacaaacaa tcagaagata tactttacaa gcaaggatgg 1020  
 tacctattat tcgacgttaa aaagattaag aaattgctgc caaactggaa ctttcactct 1080  
 atccagacgg acggcgaagg cgtctctgta ctattttcca gagaagtgga agaagtagag 1140  
 actgtttcca agaaaagtaa gaaaaataaa aaacctagag gagatgagga taggagaaat 1200  
 taccgcacca ctaatgccaa gtacgtagtg ggtgtagatc cgggaagaac taatgtcgtt 1260  
 tcctgttcgg tatttgatac ccgtcaaaaa agggtagtga gaaaacacag aatgactgcg 1320  
 aaacaatact atcaagaatc ttggatgaca gatagaagaa aggcacacga aacgtacaag 1380  
 aagaacaata aagagtacaa agaggcgtaa gaggaataaa ctaggtacga taatggcgaa 1440  
 gaaattataa atgatggtaa cggtgatact tctacacca ctaaaaaatt cgaagcttac 1500  
 ttgaaggtag tgaacgagca ctacaggtta ctgtggaacg aaaagggaaa gaaaaagtac 1560  
 aggaaaaatg ccatgaaagt atactctaga aaacaaaagt gcatacttaa ctttatagat 1620  
 gaattaatcc ctaaaaggga taaaattgaa gattaccaca ttgcttttgg ggatgcgaaa 1680  
 tttgcctgca cgggaagagg tgagcaatac gcacacctg ccaggatttt cgccaagaag 1740  
 ataaaggaaa gagtcggagg tgataagagg tttactttcg tggacgagaa atatacgtca 1800  
 aaagtatgcc atcgttgcaa tcaaccttta aatatgctgg agaaggattg tttttaccg 1860  
 aataagaaaa gaaaaccgcc gacaatagta acaaccacaa caacaacaac aacagaagaa 1920  
 gacgaagaaa atggaaaatg gaagaaggct acacctctca gagaaaatag agataccaga 1980  
 agatgctcgt ccgaaaagac gcaattcggg tacagttcaa accgaaaagt atcgacagga 2040  
 gatatctcta tggaaacgcc agtaccttct tccacttcct cttccttttg tactcctact 2100  
 tccattacat gtgtcttggg aggaaaattc gtcgacaggg acttcaatgc aagcaccaat 2160  
 attgttcata aatttctagg gttttgggat aaaaagttaa tggaaaagaa agacaagatg 2220  
 ccgttgaaat atcactttat tcgagttgcc tga 2253

<210> 119  
 <211> 746  
 <212> PRT  
 <213> SHRIMP

<400> 119  
 Met Glu Lys Lys Thr Glu Thr Ala Ala Thr Thr Glu Lys Asp Pro Glu  
 1 5 10 15  
 Pro Ser Val Ser Lys Arg Ser Arg Asn Lys Glu Pro Lys Thr Thr Ser  
 20 25 30  
 Thr Val Tyr Thr Ser Val Lys Cys Tyr Leu Ser Ser Ile Ile Lys Ser  
 35 40 45  
 Glu Ser Ser Arg Ser Asn Val Thr Ser Thr Lys Glu Arg Phe Glu Glu  
 50 55 60  
 Arg Cys Lys Ser Val Ser Lys Met Met Val Lys Gly Ser Leu Phe Leu  
 65 70 75 80  
 Arg Leu Val Val Asp Glu Cys Leu Arg Arg Tyr Asn His Leu Glu Asp  
 85 90 95  
 Glu Ile Asp Lys Trp Pro Asp Met Thr Lys Asp Asn Phe Tyr Val Gln  
 100 105 110  
 Leu Leu Arg Lys Gly Leu Asp Lys Lys Lys Leu Lys Glu Gly Ser Thr  
 115 120 125  
 His Pro Val Val Glu Asp Val Trp Asn Ser Pro Ile Val Gln Glu Thr  
 130 135 140  
 Phe Leu Ser Gln Gln Gly Glu Gly Asn Asn Pro Ile Lys Arg His Leu  
 145 150 155 160  
 Met Asp Phe Asn Thr Ile Thr Tyr Ala Ala Lys Gln Leu Lys Thr Cys  
 165 170 175  
 Phe Glu Thr Asn Leu Arg Thr His Phe Arg Thr Arg Gln Gln Arg Ala  
 180 185 190  
 Ile Ser Gly Trp Leu Ala Glu Asn Gly Phe Asp Lys Lys Tyr Thr Lys  
 195 200 205  
 Leu Val Gln His Trp Ile Ile Gly Cys Thr Tyr Lys Ser Asp Trp Val  
 210 215 220  
 Asp Ser Gly Asp Leu Glu Arg Val Lys Glu Gly Thr Lys Asn Phe Val  
 225 230 235 240

Thr	Leu	His	Arg	Lys 245	His	Leu	Cys	Val	Ile 250	Ser	Asp	Lys	Lys	Asn 255	Gly
Thr	Ile	Ser	Tyr 260	Ser	Pro	Glu	Glu	Lys 265	Tyr	Pro	Ile	Pro	Ser	Ile	Leu
Asn	Tyr	Tyr 275	Lys	Phe	Leu	Gln	Thr	Glu 280	Tyr	Pro	Gln	Asn	Lys	Lys	Ile
Gln	Lys	Met	Ile	Val	Val	Pro 295	Lys	His	Lys	Leu	Lys 300	Ile	His	Tyr	Cys
Thr 305	Phe	Asp	Gln	Thr	Thr 310	Ile	Gln	Gly	Ile	Cys 315	Lys	Asp	Leu	Gly	Val
Trp	Lys	Asp	Met	Glu 325	Glu	Arg	His	Lys	Gln 330	Ser	Glu	Asp	Ile	Leu	Tyr
Lys	Gln	Gly	Trp 340	Tyr	Leu	Leu	Phe	Asp 345	Val	Lys	Lys	Ile	Lys	Lys	Leu
Arg	Pro	Asn 355	Trp	Asn	Phe	His	Ser 360	Ile	Gln	Thr	Asp	Gly 365	Glu	Gly	Val
Ser	Val	Leu	Phe	Ser	Arg	Glu 375	Val	Glu	Glu	Val	Glu 380	Thr	Val	Ser	Lys
Lys 385	Ser	Lys	Lys	Asn	Lys	Lys	Pro	Arg	Gly	Asp 395	Glu	Asp	Arg	Arg	Asn
Tyr	Pro	Pro	Thr	Asn 405	Ala	Lys	Tyr	Val	Val	Gly	Val	Asp	Pro	Gly	Arg
Thr	Asn	Val	Val 420	Ser	Cys	Ser	Val	Phe 425	Asp	Thr	Arg	Gln	Lys	Arg	Val
Val	Arg	Lys 435	His	Arg	Met	Thr	Ala 440	Lys	Gln	Tyr	Tyr	Gln	Glu	Ser	Trp
Met	Thr 450	Asp	Arg	Arg	Lys	Ala 455	Asn	Glu	Thr	Tyr	Lys 460	Lys	Asn	Asn	Lys
Glu 465	Tyr	Lys	Glu	Ala	Leu	Glu	Glu	Ile	Thr	Arg 475	Tyr	Asp	Asn	Gly	Glu
Glu	Ile	Ile	Asn	Asp 485	Gly	Asn	Gly	Asp	Thr 490	Ser	Thr	Pro	Thr	Lys	Lys
Phe	Glu	Ala	Tyr 500	Leu	Lys	Val	Val	Asn 505	Glu	His	Tyr	Arg	Leu	Leu	Trp
Asn	Glu	Lys 515	Gly	Lys	Lys	Lys	Tyr 520	Arg	Lys	Asn	Ala	Met 525	Lys	Val	Tyr
Ser	Arg 530	Lys	Gln	Lys	Cys	Ile 535	Ser	Asn	Phe	Ile	Asp 540	Glu	Leu	Ile	Pro
Lys 545	Arg	Asp	Lys	Ile	Glu	Asp 550	Tyr	His	Ile	Ala 555	Phe	Gly	Asp	Ala	Lys
Phe	Ala	Cys	Thr	Gly 565	Arg	Gly	Glu	Gln	Tyr 570	Asp	Ala	Arg	Ile	Phe	Ala
Lys	Lys	Ile	Lys 580	Glu	Arg	Val	Gly	Gly 585	Asp	Lys	Arg	Phe	Thr	Phe	Val
Asp	Glu	Lys 595	Tyr	Thr	Ser	Lys	Val 600	Cys	His	Arg	Cys	Asn 605	Gln	Pro	Leu
Asn	Met 610	Leu	Glu	Lys	Asp	Cys 615	Phe	Ser	Pro	Asn	Lys 620	Lys	Arg	Lys	Pro
Pro 625	Thr	Ile	Val	Thr	Thr 630	Thr	Thr	Thr	Thr	Thr 635	Thr	Glu	Glu	Asp	Glu
Glu	Asn	Gly	Lys	Trp 645	Lys	Lys	Ala	Thr	Pro 650	Leu	Arg	Glu	Asn	Arg	Asp
Thr	Arg	Arg	Cys 660	Ser	Ser	Glu	Lys	Thr 665	Gln	Phe	Gly	Tyr	Ser	Ser	Asn
Arg	Lys	Val 675	Ser	Thr	Gly	Asp	Ile 680	Ser	Met	Pro	Val	Pro 685	Ser	Ser	Thr
Ser	Ser 690	Ser	Phe	Cys	Thr	Pro 695	Thr	Ser	Ile	Thr	Cys 700	Val	Leu	Gly	Gly
Lys 705	Phe	Val	Asp	Arg	Asp 710	Phe	Asn	Ala	Ser	Thr	Asn 715	Ile	Val	His	Lys
Phe	Leu	Gly	Phe	Trp	Asp	Lys	Lys	Leu	Met	Glu	Lys	Lys	Asp	Lys	Met

Pro Leu Lys Tyr His Phe Ile Arg Val Ala  
 740 745 735

<210> 120  
 <211> 411  
 <212> DNA  
 <213> SHRIMP

<400> 120  
 atggggaata gtgaatctcg ttctagtggg attgaaattg tacataaaaa tgggtgcacca 60  
 aaaagggtctc ataaaactct ttacttgtca aatagaactg aaagacatgc ccagatacaa 120  
 aagcagatcg aagagttaca tcacaaaaca aataagcaat ttgaacaggc tcaaaagggtc 180  
 ctcgacaaaa acgaagagcg aaagaagcat cagcaacagc aacaaataat aattcctcta 240  
 gaccagaaag aaagacgtgc aatattggct gaaatcgata aacacatgaa agagattgat 300  
 ggtttcatcg aggaaagtga acgtctaggt ttacttgtag atgcagaaat caataacttg 360  
 gaagaaaagg aggttgaaga ggaacatctt ttgaaacaaa aagaagacta a 411

<210> 121  
 <211> 134  
 <212> PRT  
 <213> SHRIMP

<400> 121  
 Met Gly Asn Ser Glu Ser Arg Ser Ser Gly Ile Glu Ile Val His Lys  
 1 5 10 15  
 Asn Gly Ala Pro Lys Arg Ser His Lys Thr Leu Tyr Leu Ser Asn Arg  
 20 25 30  
 Thr Glu Arg His Ala Gln Ile Gln Lys Gln Ile Glu Glu Leu His His  
 35 40 45  
 Lys Thr Asn Lys Gln Phe Glu Gln Ala Gln Lys Val Leu Asp Lys Asn  
 50 55 60  
 Glu Glu Arg Lys Lys His Gln Gln Gln Gln Gln Ile Ile Ile Pro Leu  
 65 70 75 80  
 Asp Pro Glu Glu Arg Arg Ala Ile Leu Ala Glu Ile Asp Lys His Met  
 85 90 95  
 Lys Glu Ile Asp Gly Phe Ile Glu Glu Ser Leu Gly Leu Leu Val Asp  
 100 105 110  
 Ala Glu Ile Asn Asn Leu Glu Glu Lys Glu Val Glu Glu Glu His Leu  
 115 120 125  
 Leu Lys Gln Lys Glu Asp  
 130

<210> 122  
 <211> 2772  
 <212> DNA  
 <213> SHRIMP

<400> 122  
 atgtcctcct cctcgctcgtc gtcgttctcg ttccgcctct ccacctacca gacctttctc 60  
 aaggctcttg ctcacccaga tctggtggac aagattaccc agaaatgtga cgagacagga 120  
 agaaaccaga agtgcccat ccagtttctg gccgacatct cgcacctgat ccaaggagaa 180  
 agaaatggag gaaatctgtt ccctttgcac ccgttcaaga accaaccaca tctggaacca 240  
 agaatagtgga gaagtcttca cgggagaaca ttggacaatg acattgaaga atcactactgt 300  
 tattttgtca aggatctgta taatggagta ttttctctatg tgaacggcgt caaggagtta 360  
 cagggcgttc tggacaagaa aatatctgga tctggatctg gagaatcctc ctcttctaga 420  
 gctccctga tccaataac tgaatgtgat ttgtgtaca ttttcggtac tttggtagtt 480  
 cttccccca gatctaaagc gtaccgagtc atcaactgaag ctgttctagc actccccctc 540  
 aatgaattca gtaacaactg gcctcctaca aatatcaaag gagcatacgt gtctagagat 600

```

ttcaggatgt ttaatctggt ggccgggttta gatcatatag aaggagaagt tgggggagaa 660
agtgaatggg aatccatata cgcatctgtc gtcaagcgaa tggtcacat tatgcgcaac 720
aaagctgaga agaaacctcc atcaacatct agaattttta gagtgatagt ggctgaacca 780
gttaatgatg cagtgcacaa gatccctata cgtgtactca gtaaattatt cggttcaaga 840
ctcgccggta ttctccagaa agtgactctt tattcaatgc taaatcttcc atatctcctg 900
tcttcaaatt ctatagacat caagcaagga gtaaaaggaa ttacattatc tataccatct 960
gcaagaaaat tgggattcta ttactccaa aaggatacaa cattacaatc ttccctatca 1020
caagatgttg ccgactgcat agtttcaatc aacgctggta ttattggtga tgatttctct 1080
gaaaaaatac gacagtgcac tgaggagaaa aacaagccag aaaactgttg tatgtgcttt 1140
tgtgaaattg acaagacgcc cgatttttct tatagtgaac atgtggcaag gcacaatttc 1200
ttccccgtcc acgcattctc ctcatcacat gatgacaagt gttgtggagc aaagatttgt 1260
tccgaatgta tattcccgta catcatttcc ctgtatgaga aaatgactgg tgtggcaggt 1320
gtaaaagtgt tagatttgtt ccagtgccct ggttgtaaaa gtggcatgct caatctaaag 1380
ggaagatgtt acgagttttc aaatttgtgt aagagaatga tactaccata cacatcgact 1440
cattgttctt ctctctttga tgctaccata aatcgcgagc aggtctgttt ttactcccta 1500
gagtttctcc agtatgattt tgaactcgcg aggagaattg cgcatggagc taaagacatt 1560
ccccatgttt acaataaggc agtaagaat gtaaaagatt tggatagact gtgtgccttg 1620
tactgttaca aatgcgtatc tcctgtcgta tgtgatgagc cgaatgaaag tacggactat 1680
gaaatggtgg atgtaactcc cctctcaatt aatcttaccg agattgttga ttcggaagag 1740
tatgatgatg gtcccggaaa tcatatgtgg ccagcaaat ttacctgtaa ctttattgag 1800
ggttccagtg gagaaaacacc caccattagt acgtgcagag atgctgtaac tttctagga 1860
agagcaccga ggaagaaaat ggcaggatgg gatgatcaat cggcagtggt gcaagccatt 1920
atagcgctag ccaactggag aaagagtggg gaattgccc aaaaatattgt tgatttacta 1980
gaaggggtaa atgccgtact ttatagaggc gacagtttct tgttacgtgc gataaactac 2040
ccctgtgtta ttggtagatc catgagccct agtctggaac tcgttaaaaag aaagtgtaat 2100
aaaattgtct taataaaggc cttcttccac gagaaaaggg tgcgtccaga cgcattctaa 2160
aagttacttg aatgggcaga actattagtc aaaagtatc tcatggaagt ttacttcag 2220
acgccagaat gtgtcatata ccgcgcccat tcattttagt gcaaaaactc cctcattact 2280
gacgaattgg ttacatgagc tccagatgat gccacaagaa acgcctatat ccagaacct 2340
aatgcggcca gacagaatgc ggctgctgca gcctcatatt ctgggtcgct ccccaaacct 2400
gaatttgctc cctgcaaaga aaggacgatt gaatggatgt atgaaaagga caatgatgat 2460
gttagagtgt taaattgtcc ttcatgtaaa aaggctatcc agaaatatgg aggttgtgtg 2520
aatgtgtttt gtgaatgtgg aacaaacatg tgctggatat gtgaagagaa ggtttctcct 2580
gctgattcta atcatttgtt ggagaaacac aggattgttt atagtaactg tgttaggggt 2640
aaatatgcct tagaaagtat gtacgggttt gagatttgta ccatgaaaaa tgtagaagaa 2700
ggagttaaaa attattatgt aatggagaat ggatttttct ttgatgtaca agaaatggtt 2760
gctaagaaat aa
2772

```

<210> 123  
 <211> 919  
 <212> PRT  
 <213> SHRIMP

```

<400> 123
Met Ser Ser Ser Ser Ser Ser Phe Ser Phe Arg Ile Ser Thr Tyr
1 5 10 15
Gln Thr Phe Leu Lys Ala His Pro Asp Leu Val Asp Lys Ile Thr Gln
20 25 30
Lys Cys Asp Glu Thr Gly Arg Asn Gln Lys Cys Pro Ile Gln Phe Leu
35 40 45
Ala Asp Ile Ser His Leu Ile Gln Gly Glu Arg Asn Gly Gly Asn Leu
50 55 60
Phe Pro Leu His Pro Phe Lys Asn Gln Pro His Leu Glu Pro Arg Ile
65 70 75 80
Val Gly Ser Leu His Gly Arg Thr Leu Asp Asn Asp Ile Glu Glu Ser
85 90 95
Tyr Cys Tyr Phe Val Lys Asp Leu Tyr Asn Gly Val Phe Ser Tyr Val
100 105 110
Asn Gly Val Lys Glu Leu Gln Gly Val Leu Asp Lys Lys Ile Ser Gly
115 120 125
Ser Gly Ser Gly Glu Ser Ser Ser Ser Arg Ala Pro Leu Ile Pro Ile

```

WO 01/38351

203

PCT/US00/28888

130	135	140
Thr Asp Val Asp Leu Leu Tyr Ile Phe Gly Thr Leu Val Val Leu Pro		
145	150	155
Pro Arg Ser Lys Ala Tyr Arg Val Ile Thr Glu Ala Val Leu Ala Leu		
	165	170
Pro Phe Asn Glu Phe Ser Asn Asn Trp Pro Pro Thr Asn Ile Lys Gly		
	180	185
Ala Tyr Val Ser Arg Asp Phe Arg Met Phe Asn Leu Leu Ala Gly Leu		
	195	200
Asp His Ile Glu Gly Glu Val Gly Gly Glu Ser Glu Trp Glu Ser Ile		
	210	215
His Ala Ser Val Val Lys Arg Met Val Thr Ile Met Arg Asn Lys Ala		
225	230	235
Glu Lys Lys Pro Pro Ser Thr Ser Arg Ile Phe Arg Val Tyr Val Ala		
	245	250
Glu Pro Val Asn Asp Ala Val Thr Lys Ile Pro Ile Arg Val Leu Ser		
	260	265
Lys Leu Phe Gly Ser Arg Leu Ala Gly Ile Leu Gln Lys Val Tyr Ser		
	275	280
Tyr Ser Met Leu Asn Leu Pro Tyr Leu Leu Ser Ser Asn Ser Ile Asp		
	290	295
Ile Lys Gln Gly Val Lys Gly Ile Thr Leu Ser Ile Pro Ser Ala Arg		
305	310	315
Lys Leu Gly Phe Tyr Leu Leu Gln Lys Asp Thr Thr Leu Gln Ser Ser		
	325	330
Leu Ser Gln Asp Val Ala Asp Cys Ile Val Ser Ile Asn Ala Gly Ile		
	340	345
Ile Gly Asp Asp Phe Ser Glu Lys Ile Arg Gln Cys Ile Glu Glu Lys		
	355	360
Asn Lys Pro Glu Asn Cys Cys Met Cys Phe Cys Glu Ile Asp Lys Thr		
	370	375
Pro Asp Phe Ser Tyr Ser Glu His Val Ala Arg His Asn Phe Phe Pro		
385	390	395
Val His Ala Phe Ser Ser Ser His Asp Asp Lys Cys Cys Gly Ala Lys		
	405	410
Ile Cys Ser Glu Cys Ile Phe Pro Tyr Ile Ile Ser Leu Tyr Glu Lys		
	420	425
Met Thr Gly Val Ala Gly Val Lys Val Val Asp Leu Phe Gln Cys Pro		
	435	440
Gly Cys Lys Ser Gly Met Leu Asn Leu Lys Gly Arg Cys Tyr Glu Phe		
	450	455
Ser Asn Leu Cys Lys Arg Met Ile Leu Pro Tyr Thr Ser Thr His Cys		
465	470	475
Ser Ser Leu Phe Asp Ala Thr Ile Asn Arg Ala Glu Ala Cys Phe Tyr		
	485	490
Ser Leu Glu Phe Leu Gln Tyr Asp Phe Glu Thr Ala Arg Arg Ile Ala		
	500	505
His Gly Ala Lys Asp Ile Pro His Val Tyr Asn Lys Val Val Lys Asn		
	515	520
Val Lys Asp Leu Asp Arg Leu Cys Ala Leu Tyr Cys Tyr Lys Cys Val		
	530	535
Ser Pro Val Val Cys Asp Glu Pro Asn Glu Ser Thr Asp Tyr Glu Met		
545	550	555
Val Asp Val Thr Pro Pro Leu Ile Asn Leu Thr Glu Ile Val Asp Ser		
	565	570
Glu Glu Tyr Asp Asp Gly Pro Gly Asn His Met Trp Pro Ala Lys Phe		
	580	585
Thr Cys Asn Phe Ile Ala Gly Ser Ser Gly Glu Thr Pro Thr Ile Ser		
	595	600
Thr Cys Arg Asp Ala Val Thr Phe Leu Gly Arg Ala Pro Arg Lys Lys		
	610	615
		620

atggctgaaa	ccgtcgccgt	tgatgaggtg	cctacctgtc	ccatctgtat	gggtgattac	60
gattcggaca	cggattgcta	taactgggtca	aatggaggga	tgccgtgttg	caggaaaagt	120
gtccaccctg	agtgtctttt	cacctggaga	tttgaaagagc	atatggtgaa	tgaaaatcac	180
ctgttatgtc	ccatgtgtag	ggcctatata	cccccttgtt	gggtctctccg	taaagtgtat	240
gaagaggtgt	acaagtatag	ctcttttca	tcatttttgt	tgtctgtcta	ctatgttaat	300
gatgaaggtg	taaaggatac	ccttaataag	atgtcaacta	tgtctagcacc	tactttcttt	360
gtccccaatg	ccaaaggtgt	taatgagaat	gaggatgttt	atatggagag	ggcttatacc	420
aagtttgagtt	tcatgcttga	aactctatct	agacaggaaa	tgcatgcatt	cagtgaagag	480
acctttgaag	ataatcatga	ggcagcttta	atgggttaaat	tcaaggatat	ccccccctat	540
gaatatgaag	tgaatggct	taagtatgta	gctccaata	ctattgacat	tactcaatgc	600
ttgagtaatg	atgatgatga	tgatgaaggt	gataataatg	tctcccttag	tttgttgtct	660
ggtgtgacat	ccttcaattt	cattgaagat	gatgaggata	ctgtagtgtt	tgcccccca	720
gaggtagatg	ataatgatga	ctcagagtca	cttctgact	tgactgttcc	ccctagaagt	780
aacaatatata	cccttgata	tattagtggt	attagtagtt	cacttttatg	tgtttaagt	840
gatgatgatg	atgatgacac	aatgtctttg	cctgacttga	atatgcctaq	tgcttctacg	900



```

tcattccgcc ctacctcacc cgcccctacc tcaacctctc ttaacattaa tgttaacctt 960
tgttttaatg ttgattcaga ctctgatgac gaagaagtaa taccttcttc gtcattcagtg 1020
aatcagccct ctacttcctc aggaagtagt agtagtagta gtaacagtag aaagaggcca 1080
aggtatgggc gtgacgagga caggatgtct aatatttcct ctgagagtaa gagactgtgt 1140
gtagatgtca agaggtatat gtgtagactc gataatattg atgaggagta taatgagatt 1200
gccaataggt atctggctga actttctgct cttagagaaa ggagacagga aactgagaat 1260
aagcttgag attgtatatc tagaggtaac ctgttccata caactgtcaa tgatgtaatt 1320
ggtaagagtt tgtgctctaa gaaattgaag gtgaagcgta agtacgcttc aaagtggagt 1380
gctaataagc agctaattgg ttcctgtctc attaagtcag cctctaataa tgctaggttg 1440
gatgacgaaa ttgcacatgt acacagttca ttgttgaatg ggtttgatac tgacccttcg 1500
gaagctgatac aaatatcttc cctgcccaat ctataa 1536

```

<210> 125  
 <211> 509  
 <212> PRT  
 <213> SHRIMP

<400> 125

Met	Ala	Glu	Thr	Val	Ala	Val	Asp	Glu	Val	Pro	Thr	Cys	Pro	Ile	Cys
1				5					10					15	
Met	Gly	Asp	Tyr	Asp	Ser	Asp	Thr	Asp	Cys	Tyr	Asn	Trp	Ser	Asn	Gly
			20					25					30		
Gly	Met	Pro	Cys	Cys	Arg	Lys	Ser	Val	His	Leu	Glu	Cys	Leu	Phe	Thr
		35				40						45			
Trp	Arg	Phe	Glu	Glu	His	Met	Val	Asn	Glu	Asn	His	Leu	Leu	Cys	Pro
	50					55					60				
Met	Cys	Arg	Ala	Tyr	Ile	Pro	Pro	Val	Trp	Phe	Phe	Arg	Lys	Val	Tyr
	65				70					75				80	
Glu	Glu	Val	Tyr	Lys	Tyr	Ala	Ser	Phe	His	Ser	Phe	Leu	Leu	Ser	Ala
			85						90					95	
Asp	Tyr	Val	Asn	Asp	Glu	Gly	Val	Lys	Asp	Thr	Leu	Asn	Lys	Met	Ser
			100					105					110		
Thr	Ile	Leu	Ala	Pro	Thr	Phe	Phe	Val	Pro	Asn	Ala	Lys	Gly	Val	Asn
		115					120						125		
Glu	Asn	Glu	Asp	Val	Tyr	Met	Glu	Arg	Ala	Tyr	Thr	Lys	Leu	Ser	Phe
	130					135					140				
Met	Leu	Glu	Thr	Leu	Ser	Arg	Gln	Glu	Met	His	Ala	Phe	Ser	Glu	Glu
	145				150					155				160	
Thr	Phe	Glu	Asp	Asn	His	Glu	Ala	Ala	Leu	Met	Gly	Lys	Phe	Lys	Asp
			165					170						175	
Ile	Pro	Pro	Tyr	Glu	Tyr	Glu	Gly	Glu	Trp	Leu	Lys	Tyr	Val	Ala	Pro
			180					185					190		
Asn	Thr	Ile	Asp	Ile	Thr	Gln	Cys	Leu	Ser	Asn	Asp	Asp	Asp	Asp	Asp
	195					200						205			
Glu	Gly	Asp	Asn	Asn	Val	Ser	Pro	Ser	Leu	Leu	Ser	Gly	Val	Thr	Ser
	210					215					220				
Phe	Asn	Phe	Ile	Glu	Asp	Asp	Glu	Asp	Thr	Val	Val	Phe	Val	Pro	Pro
	225				230					235				240	
Glu	Val	Asp	Asp	Asn	Asp	Asp	Ser	Glu	Ser	Leu	Pro	Asp	Leu	Thr	Val
			245					250						255	
Pro	Pro	Arg	Ser	Asn	Asn	Ile	Thr	Phe	Asp	Thr	Ile	Ser	Gly	Ile	Ser
			260					265					270		
Ser	Ser	Leu	Tyr	Asp	Val	Asn	Asp	Asp	Asp	Asp	Asp	Asp	Asp	Thr	Met
	275					280						285			
Ser	Leu	Pro	Asp	Leu	Asn	Met	Pro	Ser	Ala	Ser	Thr	Ser	Ser	Ala	Pro
	290					295					300				
Thr	Ser	Ser	Ala	Pro	Thr	Ser	Thr	Ser	Leu	Asn	Ile	Asn	Val	Asn	Leu
	305				310					315				320	
Cys	Phe	Asn	Val	Asp	Ser	Asp	Ser	Asp	Asp	Glu	Glu	Val	Ile	Pro	Ser
			325					330						335	
Ser	Ser	Ser	Val	Asn	Gln	Pro	Ser	Thr	Ser	Ser	Gly	Ser	Ser	Ser	Ser

			340						345					350			
Ser	Ser	Asn	Ser	Arg	Lys	Arg	Pro	Arg	Tyr	Gly	Arg	Asp	Glu	Asp	Arg		
		355					360					365					
Met	Ser	Asn	Ile	Ser	Ser	Glu	Ser	Lys	Arg	Leu	Cys	Val	Asp	Val	Lys		
	370					375					380						
Arg	Tyr	Met	Cys	Arg	Leu	Asp	Asn	Ile	Asp	Glu	Glu	Tyr	Asn	Glu	Ile		
385					390					395					400		
Ala	Asn	Arg	Tyr	Leu	Ala	Glu	Leu	Ser	Ala	Leu	Arg	Glu	Arg	Arg	Gln		
				405					410					415			
Glu	Thr	Glu	Asn	Lys	Leu	Gly	Asp	Cys	Ile	Ser	Arg	Gly	Asn	Leu	Phe		
			420					425					430				
His	Thr	Thr	Val	Asn	Asp	Val	Ile	Gly	Lys	Ser	Leu	Cys	Ser	Lys	Lys		
	435						440					445					
Leu	Lys	Val	Lys	Arg	Lys	Tyr	Ala	Ser	Lys	Trp	Ser	Ala	Asn	Lys	Gln		
	450					455					460						
Leu	Ile	Gly	Ser	Cys	Leu	Ile	Lys	Ser	Asn	Asn	Ala	Arg	Leu	Asp	Asp		
465					470					475					480		
Glu	Ile	Ala	His	Val	His	Ser	Ser	Leu	Leu	Asn	Gly	Phe	Asp	Thr	Asp		
				485				490						495			
Pro	Ser	Glu	Ala	Asp	Gln	Ile	Ser	Ser	Leu	Pro	Asn	Leu					
			500					505									

<210> 126  
 <211> 1746  
 <212> DNA  
 <213> SHRIMP

<400> 126  
 atgggggggac ccactgtaat tactactacc atcaatactg gtggagacca ccaccaccag 60  
 cagtatgttt accatcaggg gaataaaaaa cggcctgtgg aagaatataa caacaacaac 120  
 tacgcgtctg gttcaacctc cgaagccaca actgttcccc cttacaacaa caacaacaac 180  
 aacatcacta tcaagacttg ggatgacgtc atcaacctta gcatcacgcc ccctccccct 240  
 aaacgtttca agaagtctga agttgctccc tctcctccca ctactcgcac cttttcaaac 300  
 gtgtgtgcgt ccaaggtgat taggcagtgt aagaggcagt ataatgagt gattgaacgt 360  
 gattccccctt actactttta aggcatgtag aagagttgta gtcttgagga caattatgat 420  
 acctgtcaac agttgagaat tggccatagg tcaattgtta agtctagcaa gtatgtccat 480  
 gatacctggt tctatggaaa ggaccctaaa gttggcttct attggcccac ctcttcttgc 540  
 gatgaagaga tgagattttt tgacactaga cacattctta aggagttgtc tagtcgtaat 600  
 attccgtcctt cccagattat ggacataatg tatatggctg tagaggtgtt ccaattgcct 660  
 tcaagtgcct gtgagcgaat tagacaaaag actagcacgc taattaagga agtttctgac 720  
 cagtgtgaga actgggaaaa cttccgtaag actgctcgtg gttgtttgtc tgatttgggtc 780  
 gaagtgcctg aagatgtgaa ggactttaac actttcatct gtccttgga gaccttttt 840  
 gagattaaat atgggactta ttacattgtg aataggggga ctgttgtaaa gtttatgaag 900  
 gatatgaact atgaagagtt tgtttttgag tgtgttaatg gcctttctgt atacagaaag 960  
 aatattaagg ggtagttgg ggtgactggt gtgtgtcctc aggggttatg tttagagatg 1020  
 ccatttgcag gtatcagtat tgatgatgtc attaggtgtg tcaaggatag tttagatggt 1080  
 ggggagtatt atgagtcaag ggacgcacgc ttgttgatg ggggtgtcat gcttcaaagg 1140  
 atgggacggt taccagaggt aaagggggtt gatcacgtcg caccaataac agactctttc 1200  
 attgcccgaag aggttgtaag aagtatgttt gaaaaactaa aggtgaacat gccttttgtt 1260  
 ttggctgaga cttgtaatgt aattacaaga gttgcaaatg agggattat taatgtcgat 1320  
 ataaaggctg ataactttgt tatagatagc atatctggcc aacctaaaat gattgacttg 1380  
 ggactctcat accctctagg ttattgttac aacgatgaat attttaggaa cacggaagaa 1440  
 ctaatcaggc agtacattca cacacctccc gagttcttta ggggacactg tctaggtgcc 1500  
 tattcaatga cgtacagttt cagtgtaatg gcttccagta tactggaaga tgtgttgct 1560  
 tgttctaaca tggaaggccc tgcctttaat ttgatgtcaa acatgcactt tttgatgttg 1620  
 ttgcaaagcg gaacagacac tgatttctat caaatcgcc cttcaatcac agaatatgcc 1680  
 cttgccatga agcacatatt cccttttaag gggactgtaa tgaacctgtt taaagtaaag 1740  
 aatga 1746

<210> 127

<211> 575  
<212> PRT  
<213> SHRIMP

<400> 127

Met 1	Gly 2	Gly 3	Pro 4	Thr 5	Val 6	Ile 7	Thr 8	Thr 9	Thr 10	Ile 11	Asn 12	Thr 13	Gly 14	Gly 15	Asp 16
His 17	His 18	His 19	Gln 20	Gln 21	Tyr 22	Val 23	Tyr 24	His 25	Gln 26	Gly 27	Asn 28	Lys 29	Lys 30	Arg 31	Pro 32
Val 33	Glu 34	Glu 35	Tyr 36	Asn 37	Asn 38	Asn 39	Asn 40	Tyr 41	Ala 42	Ser 43	Gly 44	Ser 45	Thr 46	Ser 47	Glu 48
Ala 49	Thr 50	Thr 51	Val 52	Pro 53	Ala 54	Tyr 55	Asn 56	Asn 57	Asn 58	Asn 59	Asn 60	Asn 61	Ile 62	Thr 63	Ile 64
Lys 65	Thr 66	Trp 67	Asp 68	Asp 69	Val 70	Ile 71	Asn 72	Leu 73	Ser 74	Ile 75	Thr 76	Pro 77	Pro 78	Pro 79	Pro 80
Lys 81	Arg 82	Phe 83	Lys 84	Lys 85	Ser 86	Glu 87	Val 88	Ala 89	Pro 90	Ser 91	Pro 92	Pro 93	Thr 94	Thr 95	Arg 96
Thr 97	Phe 98	Ser 99	Asn 100	Val 101	Cys 102	Ala 103	Ser 104	Lys 105	Val 106	Ile 107	Arg 108	Gln 109	Cys 110	Lys 111	Arg 112
Gln 113	Tyr 114	Asn 115	Glu 116	Trp 117	Ile 118	Glu 119	Arg 120	Asp 121	Ser 122	Pro 123	Tyr 124	Tyr 125	Phe 126	Lys 127	Gly 128
Ile 129	Glu 130	Lys 131	Ser 132	Cys 133	Ser 134	Leu 135	Glu 136	Asp 137	Asn 138	Tyr 139	Asp 140	Thr 141	Cys 142	Gln 143	Gln 144
Leu 145	Arg 146	Ile 147	Gly 148	His 149	Arg 150	Ser 151	Ile 152	Val 153	Lys 154	Ser 155	Ser 156	Lys 157	Tyr 158	Val 159	His 160
Asp 161	Thr 162	Cys 163	Phe 164	Tyr 165	Gly 166	Lys 167	Asp 168	Pro 169	Lys 170	Val 171	Gly 172	Phe 173	Tyr 174	Trp 175	Pro 176
Thr 177	Ser 178	Ser 179	Cys 180	Asp 181	Glu 182	Glu 183	Met 184	Arg 185	Phe 186	Phe 187	Asp 188	Thr 189	Arg 190	His 191	Ile 192
Leu 193	Lys 194	Glu 195	Leu 196	Ser 197	Ser 198	Arg 199	Asn 200	Ile 201	Pro 202	Ser 203	Ser 204	Gln 205	Ile 206	Met 207	Asp 208
Ile 209	Met 210	Tyr 211	Met 212	Ala 213	Val 214	Glu 215	Val 216	Phe 217	Gln 218	Leu 219	Pro 220	Ser 221	Ser 222	Ala 223	Cys 224
Glu 225	Arg 226	Ile 227	Arg 228	Gln 229	Lys 230	Thr 231	Ser 232	Thr 233	Leu 234	Ile 235	Lys 236	Glu 237	Val 238	Ser 239	Asp 240
Gln 241	Cys 242	Glu 243	Asn 244	Trp 245	Glu 246	Asn 247	Phe 248	Arg 249	Lys 250	Thr 251	Arg 252	Cys 253	Leu 254	Ser 255	Asp 256
Leu 257	Val 258	Glu 259	Val 260	Pro 261	Glu 262	Asp 263	Val 264	Lys 265	Asp 266	Phe 267	Asn 268	Thr 269	Phe 270	Ile 271	Cys 272
Pro 273	Trp 274	Glu 275	Thr 276	Phe 277	Phe 278	Glu 279	Ile 280	Lys 281	Tyr 282	Gly 283	Val 284	Tyr 285	Tyr 286	Ile 287	Val 288
Asn 289	Arg 290	Gly 291	Thr 292	Val 293	Val 294	Lys 295	Phe 296	Met 297	Lys 298	Asp 299	Met 300	Asn 301	Tyr 302	Glu 303	Glu 304
Phe 305	Val 306	Phe 307	Glu 308	Cys 309	Val 310	Asn 311	Gly 312	Leu 313	Ser 314	Val 315	Tyr 316	Arg 317	Lys 318	Asn 319	Ile 320
Lys 321	Gly 322	Val 323	Val 324	Gly 325	Val 326	Thr 327	Gly 328	Val 329	Cys 330	Pro 331	Gln 332	Gly 333	Leu 334	Cys 335	Leu 336
Glu 337	Met 338	Pro 339	Phe 340	Ala 341	Gly 342	Ile 343	Ser 344	Ile 345	Asp 346	Asp 347	Val 348	Ile 349	Arg 350	Cys 351	Val 352
Lys 353	Asp 354	Ser 355	Leu 356	Asp 357	Gly 358	Gly 359	Glu 360	Tyr 361	Tyr 362	Glu 363					

<400> 129															
Met	Leu	Ser	Thr	Cys	Asp	Leu	Lys	His	Pro	Ser	Ser	Thr	Asp	Gly	Asn
1				5					10					15	
Val	Leu	Lys	Asn	Ile	His	Phe	Ser	Glu	Ser	Ile	Pro	Ala	Asn	Asp	Ile
			20					25					30		
Ile	Ser	Phe	Pro	Ser	Ser	Asp	Thr	Glu	Glu	Leu	Asn	Lys	Asp	Leu	Leu
		35					40					45			
Asp	Ser	Val	Arg	Asn	Gln	Ile	Lys	Phe	Gly	Phe	Asp	Pro	Ile	Thr	Glu
	50					55					60				
Thr	Leu	Lys	Asn	Cys	Ile	Thr	Thr	Gln	Thr	Leu	Leu	His	Ser	Phe	Leu

65					70					75				80
Lys	Ser	Ser	Leu	Leu	Thr	Leu	Gln	Glu	Lys	Phe	Asn	Glu	Trp	Gly
				85					90					95
Ile	Gln	Leu	Glu	Lys	Gly	Gly	Gln	Glu	Met	Ala	Leu	Cys	Ala	Ser
			100					105					110	Leu
Lys	Ile	Met	Gly	Gln	Ile	Ser	Ala	Leu	Ile	Glu	Thr	Ala	Lys	Glu
		115					120					125		Ala
Ser	Met	Asp	Asn	Lys	Lys	Lys	Asn	Asn	Asn	Ala	Cys	Ala	Asn	Cys
	130					135				140				Arg
Asp	Ser	Lys	Cys	Ser	Ala	Ser	Leu	Val	Thr	Leu	Phe	Asn	Lys	Thr
145					150					155				160
Asp	Glu	Lys	Tyr	Val	Lys	Gln	Asn	Ser	Ser	Ser	Ala	Ser	Ala	Leu
				165				170						175
Ala	Asn	Thr	Phe	Thr	Ala	Gly	Ala	Asn	Lys	Pro	Pro	Lys	Glu	Phe
			180					185					190	Ile
Thr	Lys	Asp	Asn	Ala	His	Gly	Asn	Ser	Asp	Thr	Asn	Tyr	Thr	Ala
	195						200					205		Met
Ser	Asp	Asn	Leu	Ile	Cys	Pro	Gly	Lys	Tyr	Tyr	Ser	Ser	Asp	Ile
	210					215					220			Thr
Tyr	Glu	Val	Thr	Lys	Gln	Ala	Lys	Glu	Arg	Ile	Lys	Asn	Asn	Asn
225					230					235				Lys
Lys	Met	Arg	Leu	Ala	Thr	Gly	Val	Glu	Met	Val	Met	Lys	Glu	Leu
				245				250						255
Ala	Glu	Asn	Asn	Lys	Glu	Gly	Gly	Arg	Val	Glu	Val	Glu	Val	Glu
			260					265					270	Gly
Val	Glu	Gln	Gln	Gln	Pro	Ser	Thr	Ser	Gly	Glu	Glu	Met	Gln	Met
	275						280					285		Glu
Ile	Met	Leu	Pro	Thr	Pro	Pro	Pro	Pro	Asp	Leu	Glu	Ser	Leu	Val
	290					295				300				Thr
Glu	Gly	Val	Asp	Asp	Tyr	Pro	Val	Phe	Ser	Pro	Leu	Pro	Ser	Leu
305					310					315				Leu
Ser	Pro	Met	Pro	Asp	Leu	Pro	Ser	Asn	Gly	Asn	Ser	Ala	Leu	Glu
			325					330						Asp
Gly	Gly	Pro	Phe	Ala	Pro	Ser	Ala	Asp	Ile	Val	Val	Asp	Lys	Thr
			340					345					350	Ser
Glu	Ile	Met	Gly	Arg	Thr	Pro	Gly	Ser	Glu	Trp	Val	His	Gln	Arg
	355						360					365		Asp
Arg	Asn	Ser	Lys	Met	Glu	Ile	Arg	Asn	Tyr	Gly	Arg	Ser	Gly	Ile
	370					375					380			Asn
Thr	Gly	Arg	Tyr	Arg	Arg	Asn	Asn	Thr	Val	Leu				
385					390					395				

<210> 130  
 <211> 1233  
 <212> DNA  
 <213> SHRIMP

<400> 130  
 atggaagaac atctatcctt caacaaaccc tctccagaaa atggagtagt cttctttgac 60  
 tttagtata atacttctat gtctaacatg gtagacaata tccggcacag acttcctatg 120  
 gacaagaaat tttctcccaa ggctctactc ttggcctcta cgccattcc cagtgcagaa 180  
 caactctcca ccaaagtaaa caaggcaata tttctccta gagaaactat tgttttatct 240  
 aaagcgtaa agattgttgt tactggactt tatgtagatg gggaatatgt gtagatgtgt 300  
 atttgcctct atccagaaaa acacaccttg aatgggattc tgagatatgt tgtgcatcta 360  
 aatatgatgt taatggacaa ggcagaagac gcagacgaaa ttcgttgttg acttattcct 420  
 ctagggagag ggtttaatag agaggctttc aagtttgttg acccagttat cccttggtgcc 480  
 ggctacaaca tattgaatgg ataccaccca gataatggcc accaaatctc cccttcatct 540  
 actcaaccac aggtgcaaag aagatgcgca gtcaagcaaa tgtacaaaca aatcaatggc 600  
 atgtttgaag tgggtgaaaca attttcaatc aaacacaaca atagaatctt cactataaat 660  
 caagttgatt ttaaagggtga agaaatgaaa atgttctttg ctctctacag cgaagagtta 720

```

cttcctttct attcagaaac gggaaaatta ttatctgaga aacacgtctc caaatcattt 780
tctcagttgc ctccacatgt gaccatttct gttttctatt tgcgaaatat ggaagaatac 840
aatactctga tgaaaacaga ttttgggagt tgttttgac cagctatcaa aatcgacact 900
ggggataatt ttgagttggt tgggatgaac aacaacattc tcgtatccaa agtatgtgtt 960
ggagatgatg cactcgattt gcgtagacga ataatggaac acatttcaga tgcaataggg 1020
agaaatggtt aactcgctga caacagactg aatcctcaca tcactcatgg gaagattaat 1080
gagggtgttg ttggtgaatg ggtgagcaga ttcgctccgt gcaatttcct ttgtaaaccg 1140
agggagaaaa ttgtattcgg aggtacaaaa ttcatttttg gaagggtcag taatggaaac 1200
tatgtaataa aacaaccagt tgattatgta tga 1233

```

&lt;210&gt; 131

&lt;211&gt; 404

&lt;212&gt; PRT

&lt;213&gt; SHRIMP

&lt;400&gt; 131

```

Met Glu Glu His Leu Ser Phe Asn Lys Pro Ser Pro Glu Asn Gly Val
 1          5          10          15
Val Phe Phe Asp Phe Ser Asp Asn Thr Ser Met Ser Asn Met Val Asp
          20          25          30
Asn Ile Arg His Arg Leu Pro Met Asp Lys Lys Phe Ser Ser Lys Ala
          35          40          45
Leu Leu Leu Ala Ser Thr Pro Ile Pro Ser Asp Glu Gln Leu Ser Thr
          50          55          60
Lys Val Asn Lys Ala Ile Phe Ser His Arg Glu Thr Ile Val Leu Ser
          65          70          75          80
Lys Ala Leu Lys Ile Val Val Thr Gly Val Asp Gly Glu Tyr Val Asp
          85          90          95
Asp Val Ile Cys Leu Tyr Pro Glu Lys His Thr Leu Asn Gly Ile Leu
          100          105          110
Arg Tyr Val Val His Leu Asn Met Met Leu Met Asp Lys Ala Glu Asp
          115          120          125
Ala Asp Glu Ile Arg Cys Gly Leu Ile Pro Leu Gly Arg Gly Phe Asn
          130          135          140
Arg Glu Ala Phe Lys Phe Val Asp Pro Val Ile Pro Cys Ala Gly Tyr
          145          150          155          160
Asn Ile Leu Asn Gly Tyr His Pro Asp Asn Gly His Gln Ile Ser Pro
          165          170          175
Ser Ser Thr Gln Pro Gln Val Gln Arg Arg Cys Ala Val Lys Gln Met
          180          185          190
Tyr Lys Gln Ile Asn Gly Met Phe Glu Val Val Lys Gln Phe Ser Ile
          195          200          205
Lys His Asn Asn Arg Ile Phe Thr Ile Asn Gln Val Asp Phe Lys Gly
          210          215          220
Glu Glu Met Lys Met Phe Phe Ala Lys Glu Glu Leu Leu Pro Phe Tyr
          225          230          235          240
Ser Glu Thr Gly Lys Leu Leu Ser Glu Lys His Val Ser Lys Ser Phe
          245          250          255
Ser Gln Leu Pro Pro His Val Thr Ile Ser Val Phe Tyr Leu Arg Asn
          260          265          270
Met Glu Glu Tyr Asn Thr Leu Met Lys Thr Asp Phe Gly Ser Cys Phe
          275          280          285
Ala Pro Ala Ile Lys Ile Asp Thr Gly Asp Asn Phe Glu Leu Phe Gly
          290          295          300
Met Asn Asn Asn Ile Leu Val Ser Lys Val Cys Val Gly Asp Asp Ala
          305          310          315          320
Leu Asp Leu Arg Arg Arg Ile Met Glu His Asp Ala Ile Gly Arg Asn
          325          330          335
Val Glu Leu Ala Asp Asn Arg Leu Asn Pro His Ile Thr His Gly Lys
          340          345          350
Ile Asn Glu Gly Val Val Gly Glu Trp Val Ser Arg Phe Ala Pro Cys

```

355	360	365
Asn Phe Leu Cys Lys Pro Arg Glu Glu Ile Val	Phe Gly Gly Thr Lys	
370	375	380
Phe Ile Phe Gly Arg Val Ser Asn Gly Asn Tyr	Val Ile Lys Gln Pro	
385	390	395
Val Asp Tyr Val		400

<210> 132  
 <211> 2839  
 <212> DNA  
 <213> SHRIMP

<400> 132

atggattcaa	cttctactac	aactatagaa	gcagaaaagg	cactcctaaa	agaatacgt	60
aatgaaaacc	tgacatggga	attcgtagac	agagttatac	ggcacgagaa	acttatgcag	120
agaactgaca	tgagactgaa	aacatcatcc	agaagactgt	tttcgttcat	ctctatatac	180
agtttccttc	aagatttctt	cactgctcgt	gatggagtaa	atagtgcaga	gtggtgtacc	240
caatcagccc	tgtaccacat	gttggatgga	gttgcttcca	taatatcttg	cttcagggaag	300
aggatagact	actacaacaa	gaagatggag	agattggcct	gcaccagtat	acgtgaagga	360
tactttcttg	tcgatgtgaa	gacaatagaa	tctagacacg	ttgaactact	agatcctgat	420
aagaaaatat	ggcaacgttt	atatgctgaa	aaaattgtct	cagaaaaagt	tgtcgtatgca	480
tacaatgaag	tgagcaaat	attgcccgat	gaagcaatgg	caaactacaa	ctacagaact	540
ggtttgggtg	atctttcgga	tactcttaaa	aatgcaaaaa	agcccccaac	tgattttaacg	600
atgacggatt	ttgactttta	tgaaaaatac	ataagggtcag	atattgtatt	gggaaaaatca	660
aacaagttgt	cgggtatggt	tagtgagaat	tttgaaatcc	ttccagacat	aaacattaaa	720
gtgcctcgta	gattggaaag	gtactttaac	gttgaaacga	attatagtct	ggagcacaaac	780
tttagattcc	ctagcaacca	catcagggga	ttaatatattg	cttatttcat	tggtaatatt	840
tttgagggtg	ccttttcatg	tgtacaactt	tacctgctcg	gtttcactct	ttctgcggct	900
agtgcattga	gagaaaacgt	actggatact	cccttctcaa	aattgaaaca	gtatataaag	960
aacgacaaca	aaactaaaaa	ttcatcctcc	aatgaagata	atgacgggga	agaatattac	1020
ccctgtgaac	tccaatatgc	aagaataaac	tctaattgaca	aaaatgcgtg	cagaaaaagt	1080
attgttaaag	cagtgaat	tgtagcagat	agagtgagaa	aagcatccgt	cacaatgatg	1140
aggacaccta	tcgccgaaca	cgaatcagat	ggctacatgg	cagactggct	ctcccttcaa	1200
atttctaaac	ttttgggacg	aaaggtgtct	gcttcttatg	ccttgttatt	tatagtgaac	1260
tggtgtgctc	acaagtataa	acagagtttt	acgaacgatg	ttaatggaag	tgaaaaatac	1320
gagatccttc	tgaagaaatt	gacagttgct	tgtgggctca	catacaacca	caagtgcggc	1380
atggttggtc	cagtaaatagg	ctttggttca	gggatgacaa	acagaaaatt	gagacagtat	1440
gcagttcatt	gcattgagaa	tgttataggg	tcttttatct	cttcggggaa	gaggaagaag	1500
gacatccatg	aagaccctaa	aaagttggaa	gaaatgtccc	tgatgcagct	ttctgcgcgc	1560
ctgtttaaga	acaacgacgt	gatgaaacga	gggcaagatg	ggaaggtaac	gttcgcgaat	1620
gaggacaatg	tacaagactt	cttggaagaa	ttaaagacaa	aggagtttgt	gcccaatgaa	1680
aggagaagaa	aaatccacga	agaggaatac	acaaaatctc	tacacaccaa	tctaaagatg	1740
acgttcagat	ttggcgtctg	tggttttcaa	caccctcttc	cagctagtag	tgacaagcca	1800
acccaagtgt	ctcttcagct	actaaaacaa	cgtaaacat	tcgtccagag	agaaactgca	1860
gctgcggtta	attggaccag	attacttcaa	tttttgttcc	cttctgacga	gagggataat	1920
aaaagacatc	aaaattcact	tccttggaat	aggctaggtt	ccaattttaa	tagacatttc	1980
atatctctag	cctctaaatt	cataaagaga	tctgtgcact	gtgaacgtgt	tgtcaatgat	2040
ataatttcaa	aattcaatgc	agatattctt	cctctcggtg	aagacccaga	ccattttcta	2100
atgaccaagg	ctggtttggt	catagaggat	catgagaggg	agaatatcga	taatgccatg	2160
tactcgttgt	gtggaggatt	taataaccag	acaacagagc	aaaagttaaa	cagtatacgt	2220
ctacgtatat	cagctgaggc	attaaaaaatt	gcaagaaatt	gtgttcttgc	aacaaccttt	2280
tcaaaatcat	acaacgaaga	caggccattt	ttacctcgca	cggatgaggc	aaagtttgtg	2340
cctatacctt	tatttggcgt	ggaaccttta	caccctttac	tcaattcatt	tattgataat	2400
actgcaataa	aatgtaatga	tagtgtatca	gatttctggt	tagaagagag	tgacgatatt	2460
tttaaagagg	cattagtttc	tcatacaatt	ctaacagatt	ccagtgtgta	ctcaactctt	2520
gttggaaga	atgaggatta	ttgtgataac	aataagagtg	gaaaacgcat	tggtaatata	2580
ctagtgtgta	cactgtatga	tagtgagg	agagccaact	ataatggtct	acattcagac	2640
aagcctagaa	agcatgatcc	tacaccatgg	agcagtaaaa	atactggcca	gtctggacgt	2700
agcacgactg	atctttcccc	caactctgtt	attgttctat	tagacactga	gaatgttgca	2760

gatgactatg aagatgagga ggaagattat gaagctttga aacaatcaga gagggataat 2820  
gtcatcacac taaataacg 2839

<210> 133

<211> 945

<212> PRT

<213> SHRIMP

<220>

<221> VARIANT

<222> (1)...(945)

<223> Xaa = Any Amino Acid

<400> 133

Met	Asp	Ser	Thr	Ser	Thr	Thr	Thr	Ile	Glu	Ala	Glu	Lys	Ala	Leu	Leu	1	5	10	15
Lys	Glu	Tyr	Val	Asn	Glu	Asn	Leu	Thr	Trp	Glu	Phe	Val	Asp	Arg	Val	20	25	30	
Ile	Arg	His	Glu	Lys	Leu	Met	Gln	Arg	Thr	Asp	Met	Arg	Leu	Lys	Thr	35	40	45	
Ser	Ser	Arg	Arg	Leu	Phe	Ser	Phe	Ile	Ser	Ile	Tyr	Ser	Phe	Leu	Gln	50	55	60	
Asp	Phe	Phe	Thr	Ala	Arg	Asp	Gly	Val	Asn	Ser	Asp	Glu	Trp	Cys	Thr	65	70	75	80
Gln	Ser	Ala	Leu	Tyr	His	Met	Leu	Asp	Gly	Val	Ala	Ser	Ile	Ile	Ser	85	90	95	
Cys	Phe	Arg	Lys	Arg	Ile	Asp	Tyr	Tyr	Asn	Lys	Lys	Met	Glu	Arg	Leu	100	105	110	
Ala	Cys	Thr	Ser	Ile	Arg	Glu	Gly	Tyr	Phe	Leu	Val	Asp	Val	Lys	Thr	115	120	125	
Ile	Glu	Ser	Arg	His	Val	Glu	Leu	Leu	Asp	Pro	Asp	Lys	Lys	Ile	Trp	130	135	140	
Gln	Arg	Leu	Tyr	Ala	Glu	Lys	Ile	Ala	Pro	Glu	Lys	Val	Val	Asp	Ala	145	150	155	160
Tyr	Asn	Glu	Val	Ser	Lys	Leu	Leu	Pro	Asp	Glu	Ala	Met	Ala	Asn	Tyr	165	170	175	
Asn	Tyr	Arg	Thr	Gly	Leu	Val	His	Leu	Ser	Asp	Thr	Leu	Lys	Asn	Ala	180	185	190	
Lys	Lys	Pro	Pro	Thr	Asp	Leu	Thr	Met	Thr	Asp	Phe	Asp	Phe	Tyr	Glu	195	200	205	
Lys	Tyr	Ile	Arg	Ser	Asp	Ile	Val	Leu	Gly	Lys	Ser	Asn	Lys	Leu	Ser	210	215	220	
Gly	Met	Phe	Ser	Glu	Asn	Phe	Glu	Ile	Leu	Pro	Asp	Ile	Asn	Ile	Lys	225	230	235	240
Val	Pro	Arg	Arg	Leu	Glu	Arg	Tyr	Phe	Asn	Val	Glu	Thr	Asn	Tyr	Ser	245	250	255	
Leu	Glu	His	Asn	Phe	Arg	Phe	Pro	Ser	Asn	His	Ile	Arg	Gly	Leu	Ile	260	265	270	
Phe	Ala	Tyr	Phe	Ile	Gly	Asn	Ile	Phe	Gly	Gly	Ala	Phe	Ser	Cys	Val	275	280	285	
Gln	Leu	Tyr	Leu	Leu	Gly	Phe	Thr	Leu	Ser	Ala	Ala	Ser	Ala	Cys	Arg	290	295	300	
Glu	Asn	Val	Leu	Asp	Thr	Pro	Phe	Ser	Lys	Leu	Lys	Gln	Tyr	Ile	Lys	305	310	315	320
Asn	Asp	Asn	Lys	Thr	Lys	Asn	Ser	Ser	Ser	Asn	Glu	Asp	Asn	Asp	Gly	325	330	335	
Glu	Glu	Tyr	Tyr	Pro	Cys	Glu	Leu	Gln	Tyr	Ala	Arg	Ile	Asn	Ser	Asn	340	345	350	
Asp	Lys	Asn	Ala	Cys	Arg	Lys	Ser	Ile	Val	Lys	Ala	Val	Lys	Phe	Val	355	360	365	
Ala	Asp	Arg	Val	Glu	Lys	Ala	Ser	Val	Thr	Met	Met	Arg	Thr	Pro	Ile				



	370					375					380					
Ala 385	Glu	His	Glu	Ser	Asp 390	Gly	Tyr	Met	Ala	Asp 395	Trp	Leu	Ser	Leu	Gln 400	
Ile	Ser	Lys	Leu	Leu 405	Gly	Arg	Lys	Val	Ser 410	Ala	Ser	Tyr	Ala	Leu 415	Leu	
Phe	Ile	Val	Asn 420	Trp	Val	Ala	His	Lys 425	Tyr	Lys	Gln	Ser	Phe	Thr	Asn	
Asp	Val	Asn 435	Gly	Ser	Glu	Lys	Tyr 440	Glu	Ile	Leu	Leu	Lys 445	Lys	Leu	Thr	
Val	Ala 450	Cys	Gly	Leu	Thr	Tyr	Asn 455	His	Lys	Cys	Gly 460	Met	Val	Val	Pro	
Val 465	Ile	Gly	Phe	Gly	Ser 470	Gly	Met	Thr	Asn 475	Arg	Lys	Leu	Arg	Gln	Tyr 480	
Ala	Val	His	Cys 485	Ile	Glu	Asn	Val	Ile	Gly 490	Ser	Phe	Ile	Ser	Ser	Gly 495	
Lys	Arg	Lys	Lys 500	Asp	Ile	His	Glu	Asp 505	Pro	Lys	Lys	Leu	Glu 510	Glu	Met	
Ser	Leu	Met 515	Gln	Leu	Ser	Ala	Arg 520	Leu	Phe	Lys	Asn	Asn 525	Asp	Val	Met	
Lys	Arg 530	Gly	Gln	Asp	Gly	Lys 535	Val	Thr	Phe	Ala	Asn 540	Glu	Asp	Asn	Val	
Gln 545	Asp	Phe	Leu	Glu	Glu 550	Leu	Lys	Thr	Lys 555	Glu	Phe	Val	Pro	Asn	Glu 560	
Arg	Arg	Arg	Lys 565	Ile	His	Glu	Glu	Glu	Tyr 570	Thr	Lys	Ser	Leu	His 575	Thr	
Asn	Leu	Lys	Met 580	Thr	Phe	Arg	Phe	Gly 585	Val	Cys	Gly	Phe	Gln 590	His	Pro	
Leu	Pro	Ala 595	Ser	Ser	Asp	Lys	Pro 600	Thr	Gln	Val	Ser	Leu	Gln 605	Leu	Leu	
Lys	Gln 610	Arg	Gln	Thr	Phe	Val 615	Gln	Arg	Glu	Thr	Ala 620	Ala	Ala	Val	Asn	
Trp 625	Thr	Arg	Leu	Leu	Gln 630	Phe	Leu	Phe	Pro	Ser 635	Asp	Glu	Arg	Asp	Asn 640	
Lys	Arg	His	Gln 645	Asn	Ser	Leu	Pro	Trp 650	Asn	Arg	Leu	Gly	Ser	Asn 655	Leu	
Asn	Arg	His	Phe 660	Ile	Ser	Leu	Ala	Ser 665	Lys	Phe	Ile	Lys	Arg 670	Ser	Val	
His	Cys	Glu 675	Arg	Val	Val	Asn	Asp 680	Ile	Ile	Ser	Lys	Phe 685	Asn	Ala	Asp	
Ile	Leu 690	Pro	Leu	Gly	Lys	Asp 695	Pro	Asp	His	Phe	Leu 700	Met	Thr	Lys	Ala	
Gly 705	Leu	Val	Ile	Glu	Asp 710	His	Ala	Arg	Glu	Asn 715	Ile	Asp	Asn	Ala	Met 720	
Tyr	Ser	Leu	Cys 725	Gly	Phe	Asn	Asn	Gln 730	Thr	Thr	Glu	Gln	Lys 735	Leu		
Asn	Ser	Ile	Arg 740	Leu	Arg	Ile	Ser	Ala 745	Glu	Ala	Leu	Lys	Asn 750	Ala	Arg	
Asn	Cys	Val 755	Leu	Ala	Thr	Thr	Phe 760	Ser	Lys	Ser	Tyr	Asn 765	Glu	Asp	Arg	
Pro	Phe 770	Leu	Pro	Arg	Thr	Asp 775	Glu	Ala	Lys	Phe	Val 780	Pro	Ile	Pro	Leu	
Phe 785	Gly	Val	Glu	Pro	Leu	His	Pro	Leu	Leu	Asn 795	Ser	Phe	Ile	Asp	Asn 800	
Thr	Ala	Asn	Lys 805	Cys	Asn	Asp	Ser	Val	Ser 810	Asp	Phe	Trp	Leu	Glu	Glu 815	
Ser	Asp	Asp	Ile 820	Phe	Lys	Glu	Ala	Leu 825	Val	Ser	His	Thr	Ile 830	Leu	Thr	
Asp	Ser	Ser 835	Val	Tyr	Ser	Thr	Leu 840	Val	Gly	Glu	Asp	Glu	Asp	Tyr	Cys	
Asp	Asn 850	Asn	Lys	Ser	Gly	Lys 855	Arg	Ile	Gly	Asn 860	Thr	Leu	Val	Cys	Thr	

Leu Tyr Asp Met Met Gly Arg Ala Asn Tyr Asn Gly Leu His Ser Asp  
 865 870 875 880  
 Lys Pro Arg Lys His Asp Pro Thr Pro Trp Ser Ser Lys Asn Thr Gly  
 885 890 895  
 Gln Ser Gly Arg Ser Thr Thr Asp Phe Ser Pro Asn Ser Val Ile Val  
 900 905 910  
 Leu Leu Asp Thr Glu Asn Val Ala Asp Asp Tyr Glu Asp Glu Glu Glu  
 915 920 925  
 Asp Tyr Glu Ala Leu Lys Gln Ser Asp Asn Val Ile Thr Leu Asn Asn  
 930 935 940  
 Xaa  
 945

<210> 134  
 <211> 4821  
 <212> DNA  
 <213> SHRIMP

<400> 134  
 atgtctcgaa actcttttacg cgtcaaagga ttaaaagaaa atgggggaat tattcccaac 60  
 cctttcgacc ccttatatgt cgatacggac gcccccttcg gtatggctgg agtcaagtc 120  
 gacatcatag ggaaagggtt tgtagaatct ctattaccag gggaaattag ttccattat 180  
 aacacatttg actgttttaa aactcctaaa aaatgtcgcg taggaggcaa cgattttgaa 240  
 tgcattagct gtcgttctct cgggggagga acttgcgtca aatccagtag agaattgaaa 300  
 actgaatacg gtatagaaga tgatgatgaa tatgacggag tttgtgtccc tctcgccgat 360  
 actatttttt cagcctcttc tgcgttcgac aagcacgacg atgatgttgc tacagatgca 420  
 gcctacagaa acgtcaaccc ttttactact gtcgaggagg cttatttgca ctacgaatct 480  
 ggaggagtca taactggagg aggaaaaaag ggaacgactt atatcaccaa aaaaaggagg 540  
 tgtgtagatt cgtctgtggt gcgaaaggat ccttcccttc tcaataaaga tccaagatta 600  
 gagcccatte ttgggtgcac agatattgta ctgtgtggag gaaaaggagt ggggcgacct 660  
 atacacccta caacattctc aattattgat gatgtggacg atatcgattt cgacataagc 720  
 agtatgactt ctactatgga ttgtttgtgc gaacctggct actcccagca gagagacct 780  
 gccaccaacg ccccaaaatg tgagaagaaa gaaggtggta tacaggaaaa ggaacaaggt 840  
 ctaggctgcc ctgtcatggt caggatagg gttgtaggag atactggaac aaagggtgt 900  
 ctttgtgatg aaagtacaca aatcagactg gaagaagtgg ccggtataga tcttcccgat 960  
 gcggcaaaaa cagattatgc tcaacccttt gtagagggtg caaaattatt gctgcaaata 1020  
 acagaaaggt acgaaacttt aggtggtagt actaaggacg cgtgcctccc tagaccagga 1080  
 aacgatacga ggatgtcggc tctcggttat tcttatgccg cttctctttt tggagagca 1140  
 cctgaaatta cagcgtttta cggaggccat ctgattacgg gaggattatt gcgcgaatct 1200  
 gccatggacg ccgcccggca ttggcactca agaattgaag attctgacga acaaggaaaa 1260  
 ctcacagtca gtgaatctgt aggaggtgta gtgccctaca gtggaacagg aagtgtggcc 1320  
 gctcacatct ggaacggaga tgcactaaac gataatgggc tggttggagc tggaggggga 1380  
 aattttactg aacatccaaa cgcacgcgtc agggctcgtc ctttacccca tagcaatat 1440  
 cccggtttag gaatcgatag catagatcat gctgtgggca taatagcgtc ccagggtaaa 1500  
 atcttccccg agacggttca catgagagca ggcgatcctt caggagtaaa gactgacaga 1560  
 agggatgcgc acaacgatac gaccatagaa acatccttcc ttaaagattc tgacaaggca 1620  
 ggttacgatt cgtacaaaaga taacccttta cagaaaactga gaaaatctca cgactctggc 1680  
 atatgcgcaa cagcatacgt cgttcccttca ctccatagag taataaagga aaaaccttct 1740  
 gctaaaaaat ataaaacagt aaataagata ttgcccttgg tacactatag acctaccgct 1800  
 aagcgaatgg ccacactccc catcgagaca attttcaagc actctctttt gacggcacia 1860  
 gaaagagacc aatctttttg gaacagtaca ttaaattcta tgatggttac aaattctagt 1920  
 aatagttttg acgagcttac aaatttactt ctggactatt ttttcccaa cttgaacggc 1980  
 gaaggaagag agaggctctg cctgcccatt caataaagat caatatacaa tgaaccacaa 2040  
 aatgccaaat tcaagaaaat tggaggtatt attctgcaac ctgttacagc gcagggtgct 2100  
 aagaaaagtt ccacctttgc aagattcagt gagaaaatat tgtctacaaa ttctcccaa 2160  
 ataattgatc attacaaggc tggttcatct gcagtgttta agaaagtggg ggaagaaaga 2220  
 gcatatgaaa tgtttgtctc cctccaact gcgtggcgca tagcttctaa tgaagggaca 2280  
 ttttctcggg gaagaggact aaataatgga atagaaggta cgggtatgcg cgaggccgaa 2340  
 agggctcgaa aaacgttatc aaaaaagccc gatatttttg caggagcaat cctcactgga 2400  
 gatggagttt taatgaatgg tgcattctca cctcttgtaa gacctatgga aattcctgca 2460

WO 01/38351

215

PCT/US00/28888

```

tcttctctcc ccgaacacac ctggttcgaa cgcagatcgc cgggtcaacgc gagaggagac 2520
cctggttcgg ccgacaactt gacagccatc aacaacactt atgatcgagt aacaaagggg 2580
gacatttagag ctatttctaaa ttogacaaca gacattaaaa cctccttcaa ttcgtacgct 2640
ccagcacgac ccttttctaa gcctctcgct cctcctgcgg gagtgtctgc ggctgctcaa 2700
gtacatcatc ttttgggagt tttaggggga tttcctttac ccataccttc cagttttctc 2760
attcaaaaga gtgtccaaga atcagtctcg aacggcacgg tcgggtccat gcatggaata 2820
gtccctctaa aattccatga aggggatgaa ttatggcaac aatgtgaagt caaggaaact 2880
gaaggagccc tcaatttcat cctcctcct atggctctct ttgagtcgct actcagggtt 2940
aggactcttt cctctgaaac gtttattagg ccagaattga tccctaatac gttcagagca 3000
gattgggggc tgagccccc caccggccggc cactacttga acggcgtgta ctccccgcct 3060
tgtgtcaggg aagaaaactg ccaatctttc ggatacccg gctcgggtgc cctctctcag 3120
tacctacaa tgatggtccc caaacccctg ggcccgaat cacattcttc gttatctaaa 3180
ttttctataa agagtattgt agaggaacaa acacggctac ttctgcca tttggagaa 3240
aagagtattt ttgagatgca agatccta atgtaaaaa tttcagata aatcggggaa 3300
ttgggagaaa aggagaactg taattgcacc aacggtttat tctgtccaa agtgaatgga 3360
gggggtagga acaaaaactg cctatagcg gcaacaccta gcagaggaaa taggcattca 3420
cgttttcccc tcatgactac cctccctaaa aacgatgttc atctctctgc agcactcctg 3480
agagctcaat ccggagacgc tagaatacta aatactatcg gagaaactaa aactaatgga 3540
agaaagataa acctaaaagc tgcaacagaa aatatatgg atatatcttc caatgttatg 3600
ttggccccct ataaattttg tgccatgcgt aggagtacgg cttacactcc ttattccaca 3660
agacaagaaa aggtacctgc tgctgttttg gacgaaagga agggtagctt cgataggaac 3720
gcagaattgt taggggatgt tggtatgaca gatattgtat ctaatgacat tttaatggag 3780
gattatgaac gcttgccggg cgtgcctcct gcggaggctg aaattttcca cattataagg 3840
gatgcagcaa aaacgggtca agagggtgca aaggctagaa ggatagtgga ctttttcgag 3900
tcttctcatg gagttactgc ctctacattt aacgtgggaa cattttcccc ctacgttgag 3960
ggtgtcaagg atatagttag cctgtatgca acgccttgct ttacagacat agattcgcca 4020
acaatatcgg ccgattctgc gacaataaat gaaggcgcaa gcattgaacc aactgatggg 4080
agtgaagttg tgggtggaag agttaacagc aatatggaaa tggtaggagg aagcactgcc 4140
ggtagtacaa agaagaggag attatctatt tcagattatg tggatttgga agaagatgag 4200
gaagttttca caatcaataa gcaaggaaa gctacagaaa atctccgcgt tagaacatca 4260
tcatcatcaa agtatgtaga gggaggacaa aaagacatgg tcggtttcta tgaagcgagt 4320
aaaagagtc ctagagtcag gagaagagtc cgtgtacttc ctgtattgac accctaccat 4380
ggaggattcg aatcttgccg acctacagct gctcagagtg cttgtacaag aggtgtgga 4440
ataacctatg cagatttcat gagaccttca gatcttagtg gaacaaaaac aacattggag 4500
ggggtaagag ttaaaggccc agaaccattt gatgatcttt caaccctcta ctttagaagc 4560
gtagggggac ctaatctaag gaaatttgca caccatcacc attttgata tgaaggata 4620
atgagcagat attattatc aagagaaaaa actgtttctg ttccagaggg tgatttgaag 4680
gatagattcc ctttgtttg tcaatcagat agaggacat tcccaccaa gagggatgga 4740
actatacaac ctttagcact tgtggacatg ggagtgttgc cagaaggcgc actcacaagg 4800
aggacaattt caatggagta a

```

<210> 135

<211> 1600

<212> PRT

<213> SHRIMP

<400> 135

```

Met Ser Arg Asn Ser Leu Arg Val Lys Gly Leu Lys Glu Asn Gly Gly
1           5           10          15
Ile Ile Pro Asn Pro Phe Asp Pro Leu Tyr Val Asp Thr Asp Ala Pro
20          25          30
Phe Gly Met Ala Gly Val Lys Ser Asp Ile Ile Gly Lys Gly Phe Val
35          40          45
Glu Ser Leu Leu Pro Gly Glu Ile Ser Ser His Tyr Asn Thr Phe Asp
50          55          60
Cys Phe Lys Thr Pro Lys Lys Cys Arg Val Gly Gly Asn Asp Phe Glu
65          70          75          80
Cys Ile Ser Cys Arg Ser Leu Gly Gly Gly Thr Cys Val Lys Ser Ser
85          90          95
Arg Glu Leu Lys Thr Glu Tyr Gly Ile Glu Asp Asp Asp Glu Tyr Asp
100         105         110

```

Gly Val Cys Val Pro Leu Ala Asp Thr Ile Phe Ser Ala Ser Ser Ala  
 115 120 125  
 Phe Asp Lys His Asp Asp Asp Val Ala Thr Asp Ala Ala Tyr Arg Asn  
 130 135 140  
 Val Asn Pro Phe Thr Thr Val Glu Glu Ala Tyr Leu His Tyr Glu Ser  
 145 150 155 160  
 Gly Gly Val Ile Thr Gly Gly Gly Lys Lys Gly Thr Thr Tyr Ile Thr  
 165 170 175  
 Lys Lys Arg Gly Cys Val Asp Ser Ser Val Val Arg Lys Asp Pro Ser  
 180 185 190  
 Leu Leu Asn Lys Asp Pro Arg Leu Glu Pro Ile Leu Gly Cys Thr Asp  
 195 200 205  
 Ile Val Leu Cys Gly Gly Lys Gly Val Gly Arg Pro Ile His Pro Thr  
 210 215 220  
 Thr Phe Ser Ile Ile Asp Asp Val Asp Asp Ile Asp Phe Asp Ile Ser  
 225 230 235 240  
 Ser Met Thr Ser Thr Met Asp Cys Leu Cys Glu Pro Gly Tyr Ser Gln  
 245 250 255  
 Gln Arg Asp Pro Ala Thr Asn Ala Pro Lys Cys Glu Lys Lys Glu Gly  
 260 265 270  
 Gly Ile Gln Glu Lys Glu Gln Gly Leu Gly Cys Pro Val Met Phe Arg  
 275 280 285  
 Tyr Gly Val Val Gly Asp Thr Gly Thr Lys Gly Cys Leu Cys Asp Glu  
 290 295 300  
 Ser Thr Gln Ile Arg Leu Glu Glu Val Ala Gly Ile Asp Leu Pro Asp  
 305 310 315 320  
 Ala Ala Lys Thr Asp Tyr Ala Gln Pro Phe Val Glu Gly Ala Lys Leu  
 325 330 335  
 Leu Leu Gln Ile Thr Glu Arg Tyr Glu Thr Leu Gly Gly Ser Thr Lys  
 340 345 350  
 Asp Ala Cys Leu Pro Arg Pro Gly Asn Asp Thr Arg Met Ser Ala Leu  
 355 360 365  
 Gly Tyr Ser Tyr Ala Ala Ser Leu Phe Gly Arg Ala Pro Glu Ile Thr  
 370 375 380  
 Ala Phe Asn Gly Gly His Leu Ile Thr Gly Gly Leu Leu Arg Glu Ser  
 385 390 395 400  
 Ala Met Asp Ala Ala Gly Asn Trp His Ser Arg Ile Glu Asp Ser Asp  
 405 410 415  
 Glu Gln Gly Lys Leu Thr Val Ser Glu Ser Val Gly Gly Val Val Pro  
 420 425 430  
 Tyr Ser Gly Thr Gly Ser Val Ala Ala His Ile Trp Asn Gly Asp Ala  
 435 440 445  
 Leu Asn Asp Asn Gly Leu Val Gly Ala Gly Gly Gly Asn Phe Thr Glu  
 450 455 460  
 His Pro Asn Ala Ser Leu Arg Val Val Pro Leu Pro His Ser Asn Ile  
 465 470 475 480  
 Pro Gly Leu Gly Ile Asp Ser Ile Asp His Ala Val Gly Ile Ile Ala  
 485 490 495  
 Ser Gln Gly Lys Ile Phe Pro Glu Thr Val His Met Arg Ala Gly Asp  
 500 505 510  
 Pro Ser Gly Val Lys Thr Asp Arg Arg Asp Ala His Asn Asp Thr Thr  
 515 520 525  
 Ile Glu Thr Ser Phe Leu Lys Asp Ser Asp Lys Ala Gly Tyr Asp Ser  
 530 535 540  
 Tyr Lys Asp Asn Pro Leu Gln Lys Leu Arg Lys Ser His Asp Ser Gly  
 545 550 555 560  
 Ile Cys Ala Thr Ala Tyr Val Val Pro Ser Leu His Arg Val Ile Lys  
 565 570 575  
 Glu Lys Pro Ser Ala Lys Asn Asp Lys Thr Val Asn Lys Ile Leu Pro  
 580 585 590  
 Leu Val His Tyr Arg Pro Thr Ala Lys Arg Met Ala His Thr Pro Ile

		595					600					605				
Glu 610	Thr 610	Ile 610	Phe 610	Lys 610	His 610	Ser 615	Leu 615	Leu 615	Thr 615	Ala 615	Gln 620	Glu 620	Arg 620	Asp 620	Gln 620	
Ser 625	Phe 625	Ala 625	Asn 625	Ser 630	Thr 630	Leu 630	Asn 630	Ser 630	Met 635	Met 635	Val 635	Thr 635	Asn 635	Ser 640	Ser 640	
Asn 645	Ser 645	Phe 645	Asp 645	Asp 645	Val 645	Thr 645	Asn 645	Leu 650	Leu 650	Leu 650	Asp 650	Tyr 650	Phe 655	Phe 655	Pro 655	
Asn 660	Leu 660	Asn 660	Gly 660	Glu 660	Gly 660	Lys 665	Glu 665	Arg 665	Ser 665	Gly 665	Leu 670	Pro 670	Ile 670	Asn 670	Thr 670	
Arg 675	Ser 675	Ile 675	Tyr 675	Asn 675	Glu 675	Pro 680	Asn 680	Asn 680	Ala 680	Lys 680	Phe 685	Lys 685	Glu 685	Ile 685	Gly 685	
Gly 690	Ile 690	Ile 690	Leu 690	Gln 690	Pro 695	Val 695	Thr 695	Ala 695	Gln 695	Gly 700	Ala 700	Lys 700	Lys 700	Ser 700	Ser 700	
Thr 705	Phe 705	Ala 705	Arg 705	Phe 710	Ser 710	Glu 710	Lys 710	Ile 710	Leu 715	Ser 715	Thr 715	Asn 715	Ser 715	Pro 720	Lys 720	
Ile 725	Ile 725	Asp 725	His 725	Tyr 725	Lys 725	Ala 725	Gly 730	Ser 730	Ser 730	Ala 730	Val 730	Phe 730	Lys 735	Lys 735	Val 735	
Gly 740	Glu 740	Lys 740	Glu 740	Ala 740	Tyr 740	Glu 740	Met 745	Phe 745	Ala 745	His 745	Pro 745	Pro 750	Thr 750	Ala 750	Trp 750	
Arg 755	Ile 755	Asn 755	Glu 755	Gly 755	Thr 755	Phe 760	Phe 760	Ser 760	Gly 760	Arg 760	Gln 765	Asn 765	Gly 765	Ile 765	Glu 765	
Gly 770	Thr 770	Gly 770	Met 770	Arg 770	Glu 775	Ala 775	Glu 775	Arg 775	Val 775	Ala 780	Lys 780	Thr 780	Leu 780	Ser 780	Lys 780	
Lys 785	Pro 785	Asp 785	Ile 785	Phe 790	Ala 790	Gly 790	Ala 795	Ile 795	Leu 795	Thr 795	Gly 795	Asp 795	Gly 795	Val 800	Leu 800	
Met 805	Asn 805	Gly 805	Ala 805	Ser 805	Ser 805	Pro 810	Leu 810	Val 810	Arg 810	Pro 810	Met 810	Glu 815	Ile 815	Pro 815	Ala 815	
Ser 820	Ser 820	Leu 820	Pro 820	Glu 820	His 820	Thr 825	Trp 825	Phe 825	Glu 825	Arg 825	Arg 830	Ser 830	Pro 830	Val 830	Asn 830	
Arg 835	Asp 835	Pro 835	Gly 835	Ser 835	Ala 835	Asp 840	Asn 840	Leu 840	Thr 840	Ala 840	Ile 845	Asn 845	Asn 845	Thr 845	Tyr 845	
Asp 850	Arg 850	Val 850	Thr 850	Lys 850	Gly 855	Asp 855	Ile 855	Arg 855	Ala 855	Ile 860	Leu 860	Asn 860	Ser 860	Thr 860	Thr 860	
Asp 865	Ile 865	Lys 865	Thr 865	Ser 870	Phe 870	Asn 870	Ser 875	Tyr 875	Ala 875	Pro 875	Ala 875	Arg 875	Pro 875	Phe 875	Ser 875	
Lys 885	Pro 885	Leu 885	Ala 885	Pro 885	Pro 885	Ala 885	Gly 885	Val 885	Ser 890	Ala 890	Ala 890	Ala 890	Gln 890	Ala 895	Thr 895	
Ser 890	Phe 890	Leu 890	Gly 890	Val 890	Leu 890	Gly 890	Gly 890	Phe 890	Pro 890	Leu 890	Pro 890	Ile 890	Pro 890	Ser 890	Ser 890	
Phe 900	Leu 900	Ile 900	Gln 900	Lys 900	Ser 900	Val 900	Gln 900	Glu 900	Ser 900	Val 900	Ser 900	Asn 900	Gly 900	Thr 900	Val 900	
Gly 910	Ser 910	Met 910	His 910	Gly 910	Ile 910	Val 910	Pro 910	Leu 910	Lys 910	Phe 910	His 910	Glu 910	Gly 910	Asp 910	Glu 910	
Leu 915	Trp 915	Gln 915	Gln 915	Cys 915	Glu 915	Val 915	Lys 915	Glu 915	Thr 915	Glu 915	Gly 915	Ala 915	Leu 915	Asn 915	Phe 915	
Ile 920	Pro 920	Pro 920	Pro 920	Met 920	Ala 920	Leu 920	Phe 920	Glu 920	Ser 920	Leu 920	Leu 920	Arg 920	Val 920	Arg 920	Thr 920	
Leu 925	Ser 925	Ser 925	Glu 925	Thr 925	Phe 925	Ile 925	Arg 925	Pro 925	Glu 925	Leu 925	Ile 925	Pro 925	Asn 925	Arg 925	Phe 925	
Arg 930																

Phe	Asp	Lys	Ile	Gly	Glu	Leu	Gly	Glu	Lys	Glu	Asn	Cys	Asn	Cys	Thr
1090						1095				1100					
Asn	Gly	Leu	Phe	Cys	Pro	Lys	Val	Asn	Gly	Gly	Gly	Arg	Asn	Lys	Thr
1105					1110					1115					1120
Asp	Pro	Ile	Ala	Ala	Thr	Pro	Ser	Arg	Gly	Asn	Arg	His	Ser	Arg	Phe
				1125					1130					1135	
Pro	Leu	Met	Thr	Thr	Leu	Pro	Lys	Asn	Asp	Val	His	Leu	Ser	Ala	Ala
			1140					1145					1150		
Leu	Leu	Arg	Ala	Gln	Ser	Gly	Asp	Ala	Arg	Ile	Leu	Asn	Thr	Ile	Gly
		1155				1160					1165				
Glu	Thr	Lys	Thr	Asn	Gly	Arg	Lys	Ile	Asn	Leu	Lys	Ala	Ala	Thr	Glu
1170					1175					1180					
Asn	Ile	Trp	Asp	Ile	Ser	Ser	Asn	Val	Met	Leu	Ala	Pro	Asn	Lys	Phe
1185					1190					1195					1200
Cys	Ala	Met	Arg	Arg	Ser	Thr	Ala	Tyr	Thr	Pro	Tyr	Ser	Thr	Arg	Gln
				1205					1210					1215	
Glu	Lys	Val	Pro	Ala	Ala	Val	Leu	Asp	Glu	Arg	Lys	Gly	Thr	Phe	Asp
			1220					1225					1230		
Arg	Asn	Ala	Glu	Leu	Leu	Gly	Asp	Val	Gly	Met	Thr	Asp	Ile	Val	Ser
		1235				1240						1245			
Asn	Asp	Ile	Leu	Met	Glu	Asp	Tyr	Glu	Arg	Leu	Pro	Gly	Val	Pro	Pro
	1250				1255					1260					
Ala	Glu	Ala	Glu	Ile	Phe	His	Ile	Ile	Arg	Asp	Ala	Ala	Lys	Thr	Gly
1265					1270					1275					1280
Gln	Glu	Gly	Ala	Lys	Ala	Arg	Arg	Ile	Val	Asp	Phe	Phe	Glu	Ser	Ser
				1285					1290					1295	
His	Gly	Val	Thr	Ala	Ser	Thr	Phe	Asn	Val	Gly	Thr	Phe	Ser	Pro	Tyr
			1300					1305					1310		
Val	Glu	Gly	Val	Lys	Asp	Ile	Val	Ser	Leu	Tyr	Ala	Thr	Pro	Cys	Phe
	1315					1320						1325			
Thr	Asp	Ile	Asp	Ser	Pro	Thr	Ile	Ser	Ala	Asp	Ser	Ala	Thr	Ile	Asn
	1330				1335					1340					
Glu	Gly	Ala	Ser	Ile	Glu	Pro	Thr	Asp	Gly	Ser	Glu	Val	Val	Val	Glu
1345					1350					1355					1360
Val	Val	Asn	Ser	Asn	Met	Glu	Met	Leu	Gly	Gly	Ser	Thr	Ala	Gly	Ser
				1365					1370					1375	
Thr	Lys	Lys	Arg	Arg	Leu	Ser	Ile	Ser	Asp	Tyr	Val	Asp	Leu	Glu	Glu
			1380					1385					1390		
Asp	Ala	Glu	Val	Phe	Thr	Ile	Asn	Lys	Gln	Gly	Lys	Ala	Thr	Glu	Asn
	1395					1400						1405			
Leu	Arg	Val	Arg	Thr	Ser	Ser	Ser	Ser	Lys	Tyr	Val	Glu	Gly	Gly	Gln
	1410				1415					1420					
Lys	Asp	Met	Val	Gly	Phe	Tyr	Glu	Ala	Ser	Lys	Arg	Val	Pro	Arg	Val
1425					1430	</									

1570					1575					1580					
Gly	Val	Leu	Pro	Glu	Gly	Ala	Leu	Thr	Arg	Arg	Thr	Ile	Ser	Met	Glu
1585					1590					1595					1600

<210>	136
<211>	3483
<212>	DNA
<213>	SHRIMP

<400>	136					
atgaaaattg	tccagaacaa	ttttacgccc	gatgaaagat	ctggtgtgat	acacataaga	60
aaacctgcca	aaatagaaaa	agctgtattc	gggaatatag	cagcggctat	cgatgatccc	120
gctgctgtta	gaaaagacc	gaaaaagaaa	aggaatttga	agaattgggt	agaacggcg	180
tcaaaaaaac	tcgctaaaaa	tattgaaagg	atatcctctg	aggaattaaa	acgtgtcact	240
gatgtgcagg	atccaaaatt	gttacacagt	ataatgaaac	gcactgcacg	gcagatcggt	300
tatgatattg	gtgatgatat	ttctcctcaa	tcggctcctg	atagagatgg	ttcatcatca	360
tcatcattgc	tgccaattcg	aatgattaat	atccgtacag	aagaattatt	agaaaaaggt	420
ggaaaagata	ccatcgtgag	gatacatatt	cttgacggaa	tattaccaga	taatgtccct	480
cttcccttta	aggcagagat	aaaagtgtat	ttagtggatg	aaaaattatga	aggagaagac	540
gggggtggca	gcagcgatag	tggtcctgtc	ctattcgaaa	ctttccctga	atttgtccca	600
gcaggttggc	ctcctatcac	caatgacccg	aatgcgttct	ctcgcaacaa	tggcaataaa	660
caacaaggccg	tattcaaaca	cgtggaagta	aattctctag	ctgacggtat	aacactatca	720
acaaaaggaa	gtattttcaa	cactgggaat	agactgaaaa	tttccattgt	aacagaggat	780
aaaaacaaga	ccgtgttatt	cgattctcag	gtacaatat	cttcaccat	accaaaaaata	840
acggaagttt	ttgcttgtag	aaatgtagcc	ctcatgaggc	ttgatatgcc	taaagcaata	900
aactatgata	atgtggaata	cactcctgat	acattgaatg	aaaaatatgt	ttcagattac	960
cctgctaatt	ttccccgtct	ttcccgacaa	gctgaaatag	cgtccaattt	agcagctaaa	1020
ctacctcgcg	aaaaactaac	ttcagatat	aacaaaacct	cagtgtcgtt	cggtgactcg	1080
aaaaacaaca	ctgttaaac	accagtattt	aacaagtag	tatacaatga	aacgttaaa	1140
aatatggagg	gaaatgaatc	cgagggttac	aagattctaa	acgccacaga	gataaccac	1200
cttagaaatc	catctaacc	agctcggacc	tttatatgtg	tgtctgtacc	tgaagtgaa	1260
atagaagcgc	aatggaaaat	gttgggctgg	attgtgggct	tcaaaacgtc	gtctgacgtt	1320
ttaaccacat	cttctgggta	caatatgtta	tttccagcat	caaaagttac	tcaatctgat	1380
aaattgtttc	cggttataag	tactgatgtg	aatgctaaca	cgaacaagt	tgttgtacat	1440
aatacacctt	cgagagtcgg	gtgcttcggt	tctagcgtaa	actttaagat	agatgctgca	1500
actgcaccag	attggcctgg	acccacgaac	gggccagatt	tcttctcta	ccaactaaga	1560
ccatgtatat	tcttgaaaac	tgataatgat	aatagagaac	ctcgaataac	ggctgtgctt	1620
tcttcccccg	ccacagaata	tgcgcgtgaa	cgcacaacat	ctctactgcc	tcgtgcttta	1680
acggtttcag	ttggtccatt	gactgaagta	cgaggtggag	atatatttac	ccccgtcaa	1740
actgcattac	tgggagggtg	acagcctaca	tttaaggcac	ccgcagaacc	tactaaactg	1800
tacgtgtttt	tcctgttttt	agattcgcac	aacggcctag	tgaagcgag	tgataaccca	1860
ttccaaccca	ttcacagtat	cacgtcaagg	aacaaaacta	ctgtcttgac	cgtatctgat	1920
gttatcgtca	atgatgatga	tgatgatgtt	gtccttgaag	ataaaagcta	ccacatcaca	1980
gtctccgac	ccgtgtctgg	ttctatactt	gcaaaagaaa	acgtttcttc	ttcgagaata	2040
accagccgcc	ctatatattt	tgacggtgca	cgagacgata	gggttttcag	tgtaaagatg	2100
gaagtatttg	gagggtgacga	taaaggcata	cagatgcctt	ttacgatgga	cgggcatttt	2160
gaaggacagt	tttcagacat	gtctgtacca	tctaacgaat	tggctatatg	gaatgatccc	2220
tcaaccttta	cggcacctgt	aagggacaca	cccgccaccg	acataaccaa	caagggcata	2280
gtctactgtc	gcactactgt	tctctccatt	tctaataagag	gataaggga	ccctttcatg	2340
aaacagacta	gtcttgtccc	cctacctaca	agcataccctg	aatggcgctt	cgttgattac	2400
ggggggagaaa	taaagtaccc	tcggcacatt	ttcatttctt	ccataagaac	gaatgatact	2460
acaaaacattg	tcaatacaga	tactcaaact	gaattttcaa	tagaaaactg	gttgagggaa	2520
caaatcgata	aggaacaaga	acgtcatacga	cagttattgc	cagcacctga	tgaagcgtat	2580
acacaggggg	agaaaagtga	cgctaaaatg	tacatgggag	acggcgtttc	tgaagaacaa	2640
ttagaccaaa	ttgtgcacac	ttcaaatact	acatacgtgg	tagacgagtc	tgaaaccaaa	2700
aaggaaaaatc	ttctggtaaa	taaagaagac	aagaaacttg	cggctatttt	aggaaaatgg	2760
ggtattgttg	tttttggaag	aaataaaatc	cctgacgaac	ctgcagacag	atacacgaac	2820
tggagaaata	ctggccgcct	acgcgcagtg	ggttcttatt	ctcaatttga	acaaccagtt	2880
gcccctctac	agactcggct	agccacatgg	cttagcggcg	acctgtttac	tcggttagct	2940
gatggacaat	ttctggtacg	tctagaccct	aqatgttgaq	qaatttqtag	tqcaaacqqq	3000

WO 01/38351

220

PCT/US00/28888

ttctacaata ataatggtgc aaacaatgaa tttacttcct cgttactatt tgctatagtt 3060  
 gggaatcaag acaaggtggt gtcgtatgca gaaagagtaa gattctacat gaaaattgta 3120  
 gctagaaatg aaggaaagaa gcatttgaaa aatgacgatg gtctagtttt agtggataga 3180  
 aattcagcac tacatcgccg tctctggaac aggacgacat tcgaccatga tgatattgta 3240  
 ttatgtgtta aaatccctca aaatgtcatg tctaaaatag aacccggcac tagctcaggt 3300  
 gtactagtag accctctggt atttgccaat gtagcgtaaa gtacagacag ggaagaattc 3360  
 tacaagaaat tcatagacac ttcttcaggc cctgttggtta ttgaccgtgc atctgtttaca 3420  
 tcattcttata atatatctgt acccttaaac ttttatacca catgtggatt tatagttggt 3480  
 taa 3483

<210> 137

<211> 1156

<212> PRT

<213> SHRIMP

<400> 137

Met	Lys	Ile	Val	Gln	Asn	Asn	Phe	Thr	Pro	Asp	Glu	Arg	Ser	Gly	Val
1				5					10					15	
Ile	His	Ile	Arg	Lys	Pro	Ala	Lys	Ile	Glu	Lys	Ala	Val	Phe	Gly	Asn
			20					25					30		
Ile	Ala	Ala	Ala	Ile	Asp	Asp	Ser	Ala	Ala	Val	Arg	Lys	Asp	Pro	Lys
		35					40					45			
Lys	Lys	Arg	Asn	Leu	Lys	Asn	Gly	Leu	Glu	Pro	Ala	Ser	Lys	Lys	Leu
	50					55					60				
Ala	Lys	Asn	Ile	Glu	Arg	Ile	Ser	Ser	Glu	Glu	Leu	Lys	Arg	Val	Thr
65					70					75					80
Asp	Val	Gln	Asp	Pro	Lys	Leu	Leu	His	Ser	Ile	Met	Lys	Arg	Thr	Ala
				85					90					95	
Arg	Gln	Ile	Gly	Tyr	Asp	Ile	Gly	Asp	Asp	Ile	Ser	Pro	Gln	Ser	Ala
			100					105					110		
Pro	Asp	Arg	Asp	Gly	Ser	Ser	Ser	Ser	Ser	Leu	Leu	Pro	Ile	Arg	Met
		115					120					125			
Ile	Asn	Ile	Arg	Thr	Glu	Glu	Leu	Leu	Glu	Lys	Gly	Gly	Lys	Asp	Thr
	130						135				140				
Ile	Val	Arg	Ile	His	Ile	Leu	Asp	Gly	Ile	Leu	Pro	Asp	Asn	Val	Pro
145					150					155					160
Leu	Pro	Phe	Lys	Ala	Glu	Ile	Lys	Val	Asp	Leu	Val	Asp	Glu	Lys	Tyr
			165					170						175	
Glu	Gly	Glu	Asp	Gly	Gly	Gly	Ser	Ser	Asp	Ser	Gly	Pro	Ser	Leu	Phe
		180					185					190			
Glu	Thr	Phe	Pro	Glu	Phe	Val	Pro	Ala	Gly	Trp	Pro	Pro	Ile	Thr	Asn
	195						200					205			
Asp	Pro	Asn	Ala	Phe	Ser	Arg	Asn	Asn	Gly	Asn	Lys	Gln	Gln	Ala	Val
	210					215					220				
Phe	Lys	His	Val	Glu	Val	Asn	Ser	Leu	Ala	Asp	Gly	Ile	Thr	Leu	Ser
225					230					235					240
Thr	Lys	Gly	Ser	Ile	Phe	Asn	Thr	Gly	Asn	Arg	Leu	Lys	Ile	Ser	Ile
			245						250					255	
Val	Thr	Glu	Asp	Lys	Asn	Lys	Thr	Val	Leu	Phe	Asp	Ser	Gln	Val	Thr
		260						265					270		
Ile	Ser	Ser	Pro	Ile	Pro	Lys	Ile	Thr	Glu	Val	Phe	Ala	Cys	Arg	Asn
	275						280					285			
Val	Met	Arg	Leu	Asp	Met	Pro	Lys	Ala	Ile	Asn	Tyr	Asp	Asn	Val	Glu
	290					295					300				
Tyr	Thr	Pro	Asp	Thr	Leu	Asn	Glu	Lys	Tyr	Val	Ser	Asp	Tyr	Pro	Ala
305					310					315					320
Asn	Phe	Pro	Arg	Leu	Ser	Arg	Gln	Ala	Glu	Ile	Asn	Leu	Ala	Ala	Lys
			325						330					335	
Leu	Pro	Arg	Glu	Asn	Gln	Leu	Ser	Asp	Ile	Asn	Lys	Pro	Ser	Val	Ser
		340						345					350		
Phe	Val	Tyr	Ser	Lys	Thr	Asn	Thr	Val	Asn	Thr	Pro	Val	Leu	Asn	Lys



Val	Leu	Tyr	Asn	Glu	Thr	Leu	Lys	Asn	Met	Glu	Gly	Asn	Glu	Ser	Glu
	370					375					380				
Gly	Tyr	Lys	Ile	Leu	Asn	Ala	Thr	Glu	Ile	Thr	His	Leu	Arg	Asn	Pro
385					390					395					400
Ser	Asn	Pro	Ala	Arg	Thr	Phe	Ile	Cys	Val	Ser	Val	Pro	Glu	Ser	Glu
				405					410					415	
Ile	Glu	Ala	Gln	Trp	Lys	Met	Leu	Gly	Trp	Ile	Val	Gly	Phe	Lys	Thr
			420					425				430			
Ser	Ser	Asp	Val	Leu	Thr	Thr	Ser	Ser	Gly	Tyr	Asn	Ile	Val	Phe	Pro
		435					440					445			
Ala	Ser	Lys	Val	Thr	Gln	Ser	Asp	Lys	Leu	Phe	Ser	Val	Ile	Ser	Thr
	450					455					460				
Asp	Val	Asn	Ala	Asn	Thr	Asn	Lys	Val	Val	Val	His	Asn	Thr	Pro	Ser
465					470					475					480
Arg	Val	Gly	Cys	Phe	Gly	Ser	Ser	Val	Asn	Phe	Arg	Val	Asp	Ala	Ala
				485					490					495	
Thr	Ala	Pro	Asp	Trp	Pro	Gly	Pro	Thr	Asn	Gly	Pro	Asp	Phe	Phe	Ser
			500					505				510			
Tyr	Gln	Leu	Arg	Pro	Cys	Ile	Ile	Leu	Lys	Thr	Asp	Asn	Asp	Asn	Arg
		515					520					525			
Glu	Pro	Arg	Ile	Thr	Ala	Val	Leu	Ser	Ser	Pro	Ala	Thr	Glu	Tyr	Ala
	530					535					540				
Gly	Glu	Arg	Thr	Thr	Ser	Leu	Leu	Pro	Arg	Ala	Leu	Asn	Val	Ser	Val
545					550					555					560
Gly	Pro	Leu	Thr	Glu	Val	Arg	Gly	Gly	Asp	Ile	Ile	Thr	Pro	Val	Gln
				565					570					575	
Thr	Ala	Leu	Leu	Gly	Gly	Glu	Gln	Pro	Thr	Phe	Lys	Ala	Pro	Ala	Glu
			580					585				590			
Pro	Thr	Lys	Leu	Tyr	Ala	Val	Phe	Pro	Val	Leu	Asp	Ser	His	Asn	Gly
		595					600					605			
Leu	Val	Lys	Ala	Ser	Asp	Asn	Pro	Phe	Gln	Pro	Ile	His	Ser	Ile	Thr
	610					615					620				
Ser	Arg	Asn	Lys	Thr	Thr	Val	Leu	Thr	Val	Ser	Asp	Val	Ile	Val	Asn
625					630					635					640
Asp	Asp	Asp	Asp	Asp	Val	Val	Leu	Glu	Asp	Lys	Ser	Tyr	His	Ile	Thr
				645					650					655	
Val	Ser	Asp	Pro	Val	Ser	Gly	Ser	Ile	Leu	Ala	Lys	Glu	Asn	Val	Leu
			660					665				670			
Ser	Ser	Arg	Ile	Thr	Ser	Arg	Pro	Ile	Phe	Ile	Asp	Gly	Ala	Arg	Asp
		675					680					685			
Asp	Arg	Val	Phe	Ser	Val	Lys	Met	Glu	Val	Phe	Gly	Gly	Asp	Asp	Lys
	690					695					700				
Gly	Ile	Gln	Met	Pro	Phe	Thr	Met	Asp	Gly	His	Phe	Glu	Gly	Gln	Phe
705					710										

WO 01/38351

PCT/US00/28888

222

Leu Pro Ala Pro Ser Glu Ala Tyr Thr Gln Gly Glu Lys Val Tyr Ala  
850 855 860  
Lys Met Tyr Met Gly Asp Gly Val Ser Glu Glu Thr Leu Asp Gln Ile  
865 870 875 880  
Val His Thr Ser Asn Thr Thr Tyr Val Val Asp Glu Ser Gly Thr Lys  
885 890 895  
Lys Glu Asn Leu Leu Val Asn Lys Glu Asp Lys Lys Leu Ala Ala Ile  
900 905 910  
Leu Gly Lys Trp Gly Ile Val Val Phe Gly Ala Asn Lys Tyr Pro Asp  
915 920 925  
Glu Pro Ala Asp Arg Tyr Thr Asn Trp Arg Asn Thr Gly Arg Leu Arg  
930 935 940  
Ala Val Gly Ser Tyr Ser Gln Leu Arg Gln Pro Val Ala Pro Leu Gln  
945 950 955 960  
Thr Arg Leu Ala Thr Trp Pro Ser Gly Asp Pro Val Thr Arg Leu Ala  
965 970 975  
Asp Gly Gln Phe Leu Val Arg Leu Asp Pro Arg Cys Gly Gly Ile Gly  
980 985 990  
Ser Ala Asn Gly Phe Tyr Asn Asn Asn Gly Ala Asn Asn Glu Phe Thr  
995 1000 1005  
Ser Ser Leu Leu Phe Ala Ile Val Gly Asn Gln Asp Lys Val Val Ser  
1010 1015 1020  
Tyr Ala Glu Arg Val Arg Phe Tyr Met Lys Ile Val Ala Arg Asn Glu  
1025 1030 1035 1040  
Gly Lys Lys His Leu Lys Asn Asp Asp Gly Leu Val Leu Val Asp Arg  
1045 1050 1055  
Asn Ser Ala Leu His Arg Arg Leu Trp Asn Arg Thr Thr Phe Asp His  
1060 1065 1070  
Asp Asp Ile Val Leu Cys Val Lys Ile Pro Gln Asn Val Met Ser Lys  
1075 1080 1085  
Ile Glu Pro Gly Thr Ser Ser Gly Val Leu Val Asp Pro Leu Val Phe  
1090 1095 1100  
Ala Asn Val Ala Ser Ser Thr Asp Arg Glu Glu Phe Tyr Lys Lys Phe  
1105 1110 1115 1120  
Ile Asp Thr Ser Ser Gly Pro Val Val Ile Asp Arg Ala Ser Val Thr  
1125 1130 1135  
Ser Ser Tyr Asn Ile Ser Val Pro Leu Asn Phe Tyr Thr Thr Cys Gly  
1140 1145 1150  
Phe Ile Val Gly  
1155

<210> 138  
<211> 615  
<212> DNA  
<213> SHRIMP

<400> 138  
atggaatttg gcaacctaac aaacctggac gttgcaatta ttgcaatctt gtccattgca 60  
atcattgctc taatcggtat catgggtata atgattgtat tcaacacacg tgttggaaga 120  
agcgctcgctg ctaattatga tcagatgatg cgagtcctcaa ttcaaagaag ggcaaaggta 180  
atgtcaattc gtggagagag gtcctacaat actcctcttg gaaagggtgc catgaagaat 240  
ggtctctccg ataaggacat gaaggatgtt tctgctgac ttgtcatctc taccgtcaca 300  
gccccaaagga ctgatcccg cttggcactggg gccgagaact ctaacatgac tttgaagatc 360  
ctcaacaaca ctggcgctga tctcttgatc aacgacatta ctgttcggcc aactgttatt 420  
gcaggaaaca ttaagggaaa tactatgtcg aacacttact tctcgagcaa ggacattaaa 480  
tcttcatctt caaaaattac cctcattgac gtgtgcagca aatttgaaga cggcgagcc 540  
ttcgaagcta caatgaacat tggattcacc tccaagaatg tgatcgatat caaggacgaa 600  
atcaagaaga agtaa 615

<210> 139

<211> 204  
<212> PRT  
<213> SHRIMP

<400> 139

Met	Glu	Phe	Gly	Asn	Leu	Thr	Asn	Leu	Asp	Val	Ala	Ile	Ile	Ala	Ile
1				5					10					15	
Leu	Ser	Ile	Ala	Ile	Ile	Ala	Leu	Ile	Val	Ile	Met	Val	Ile	Met	Ile
			20					25					30		
Val	Phe	Asn	Thr	Arg	Val	Gly	Arg	Ser	Val	Val	Ala	Asn	Tyr	Asp	Gln
		35				40						45			
Met	Met	Arg	Val	Pro	Ile	Gln	Arg	Arg	Ala	Lys	Val	Met	Ser	Ile	Arg
	50					55				60					
Gly	Glu	Arg	Ser	Tyr	Asn	Thr	Pro	Leu	Gly	Lys	Val	Ala	Met	Lys	Asn
65					70					75					80
Gly	Leu	Ser	Asp	Lys	Asp	Met	Lys	Asp	Val	Ser	Ala	Asp	Leu	Val	Ile
			85						90					95	
Ser	Thr	Val	Thr	Ala	Pro	Arg	Thr	Asp	Pro	Ala	Gly	Thr	Gly	Ala	Glu
			100					105					110		
Asn	Ser	Asn	Met	Thr	Leu	Lys	Ile	Leu	Asn	Asn	Thr	Gly	Val	Asp	Leu
		115					120					125			
Leu	Ile	Asn	Asp	Ile	Thr	Val	Arg	Pro	Thr	Val	Ile	Ala	Gly	Asn	Ile
	130					135					140				
Lys	Gly	Asn	Thr	Met	Ser	Asn	Thr	Tyr	Phe	Ser	Ser	Lys	Asp	Ile	Lys
145					150					155					160
Ser	Ser	Ser	Ser	Lys	Ile	Thr	Leu	Ile	Asp	Val	Cys	Ser	Lys	Phe	Glu
				165					170					175	
Asp	Gly	Ala	Ala	Phe	Glu	Ala	Thr	Met	Asn	Ile	Gly	Phe	Thr	Ser	Lys
			180					185					190		
Asn	Val	Ile	Asp	Ile	Lys	Asp	Glu	Ile	Lys	Lys	Lys				
		195					200								

<210> 140  
<211> 2571  
<212> DNA  
<213> SHRIMP

<400> 140

atgacagagt	caaaggacta	tgttctggca	ttagtagcag	aaactaaaac	cgatgaaaaa	60
cgactaaatt	atgtatcgga	agggctagtt	gctgccattt	caaatttaca	aaatacacct	120
gaaaaacaaa	ggaaggttgt	tatttcttct	gacgtattcg	gaccacgtg	gttcaataaa	180
acgacagaat	tttttaattc	gggactaaga	ttagccaaag	gacacctatc	taaagatgcc	240
gtcatgagga	gtgtctacag	agatatagaa	ggtgttagag	agcatataat	tgacccctct	300
tggagattga	ccgaaactgc	agcgggaagaa	ctatgtgact	ttactttttt	gaaacaagca	360
ccattattga	accttttgaa	tgcctttgag	aatattatgg	acggtgtggt	tagatcagca	420
gctaattctag	ttcttttattc	cactagaggt	gataactaatg	aaccttcatg	ggtaattgat	480
tcagaaatgt	tggccaacag	aaacaattca	acggtggccg	atctggcaat	gggacgagca	540
aaaagagcca	ttgctctatt	tctaggatat	actctttgtg	atatttctaag	atggaagcaa	600
tctatcgctt	ctagaatgaa	agagaggggt	ttagatccct	ttgctgccat	gcctcctcac	660
ctagagtacg	gacgcgctgc	tgacatgatt	gaaaaaagga	taaaagactt	tattgaaggg	720
tccttttagtg	atggagttac	agtcagtgaa	gaagatggac	aatctttatgt	tgttccaact	780
atatcgactg	ttttaaccac	catggtttct	gtaatacagg	aaggattcta	ccccccaag	840
gtaggttctt	tccatgaagc	tcttttaggg	agggaaatca	tggtactatt	atcagcagcc	900
atcgacgcag	agtacagggc	agtgtgtgcc	aggaccagaa	acgctaaacc	aaatcctttg	960
acaactaaac	tggataaata	tgtgaacaat	ccccatctgc	agatgccatc	cgaaagtgtc	1020
acagaaaggg	aaaaagaatg	ggtcgagagg	gaaagagaac	gaattaaaac	tacggatatg	1080
actgctgaga	atctctttag	ggaccaccca	tatcttccaa	aggcgataga	cggaaattta	1140
gggccttaac	gcacacctac	cgcccttcaa	gtcttccaac	gtgaatacaa	gagatgtaac	1200
aagtttaacg	atattgtttc	accagaaact	ttggagtatt	ttcttgtcaa	taatagacaa	1260
gttatgtttc	ccaactactc	tgttaccgcg	gttcttgatc	ccgactcagc	tgcccgatgt	1320

```

tcaatgtatg tgctatggaa tgcattatct ttatgctctg gagggctaac acagaaaaact 1380
aatagctctg cagtgaagaa cagggttaatt cttcaagtgt ttttgaaaga tatgcattct 1440
ttattcgtct gccagcgttg tgaatctggt ttcataacaa aatctctaga cactttttacc 1500
atcagtttaa aggaacaaaag taaaccgtct atggggagaac aggaactaga gacgtactgg 1560
aaagcagtag tcgatgcttt aggtgggtggg ggtggtaata ataaaggagc agaaaatgtc 1620
aatgggttag gggagttgat gggttagata ctctctgcag attcgggtct attaagaggg 1680
ggaggggttag gtggtgatat aggtttttgaa ggaaaaatga aacaaaaaag agaggacgaa 1740
gaagtgcgta acatgcacct agtggataag aaagggtagt tgtttgaagc tgcaaaatat 1800
gtccatgtat ctaaagggtt tgctgcccct tctttctacc tgttgtacgc ggctgcagcc 1860
acctccaatc cttcaataac gaacaatttc gacagagcag tatattttact attggcaaga 1920
tggggagatt tgaagtttcc tacacataac ttatggggaa atgttccaac agatgaaaat 1980
acatcatcat tattatcctt tgcaagcttt tgggctttta ggaatgctgt tagagcacga 2040
agaaatgtga ttgatgcgtc aaacactagt tttgtaccag ggaggccctt acctctctta 2100
tctgcatttt catccaagat gttagtggat aacatgttga agaataatta tgtcaaaagt 2160
gaaaatgtta acagagaaaa attgatctgg aaagcattta gagagatgca aacggaatca 2220
gagatatgga aaactagtgg gagtaaaagc gcattccgaca ggaatgtgaa gaaggccaag 2280
caagacttga tacggaacgc atctatagga agacttattg tcgaacctgt aggaaagacc 2340
cctatttcat ccatagcctt attcaggtct atgaagagga gtagaagcga agattttaaag 2400
atgggaagca ataataagta ccgtctcgcc agggacacga aaacagctac accaaggaac 2460
cctctaagtt atactggaaa gattgtattc tctttggatg atttgaaaaa tttcagtaag 2520
gattcttata ctacaatgaa agtatttctt ctaacacctt tagatggata g 2571

```

&lt;210&gt; 141

&lt;211&gt; 852

&lt;212&gt; PRT

&lt;213&gt; SHRIMP

&lt;400&gt; 141

```

Met Thr Glu Ser Lys Asp Tyr Val Leu Ala Leu Val Ala Glu Thr Lys
 1          5          10          15
Thr Asp Glu Lys Arg Leu Asn Tyr Val Ser Glu Gly Leu Val Ala Ala
 20          25          30
Ile Ser Asn Leu Gln Asn Thr Pro Glu Lys Gln Arg Lys Val Val Ile
 35          40          45
Ser Ser Asp Val Phe Gly Pro Thr Trp Phe Asn Lys Thr Thr Glu Phe
 50          55          60
Phe Asn Ser Gly Leu Arg Leu Ala Lys Gly His Leu Ser Lys Asp Ala
 65          70          75          80
Val Met Arg Ser Val Tyr Arg Asp Ile Glu Gly Val Arg Glu His Ile
 85          90          95
Ile Asp Pro Ser Trp Arg Leu Thr Glu Thr Ala Ala Glu Glu Leu Cys
100          105          110
Asp Phe Thr Phe Leu Lys Gln Ala Pro Leu Leu Asn Leu Leu Asn Ala
115          120          125
Phe Glu Asn Ile Met Asp Gly Val Phe Arg Ser Ala Ala Asn Leu Val
130          135          140
Lys Thr Arg Gly Asp Thr Asn Glu Pro Ser Trp Val Ile Asp Ser Glu
145          150          155          160
Met Leu Ala Asn Arg Asn Asn Ser Thr Val Ala Asp Leu Ala Met Gly
165          170          175
Arg Ala Lys Arg Ala Ile Ala Leu Phe Leu Gly Tyr Thr Leu Cys Asp
180          185          190
Ile Leu Arg Trp Lys Gln Ser Ile Ala Ser Arg Met Lys Glu Arg Gly
195          200          205
Leu Asp Pro Phe Ala Ala Met Pro Pro His Leu Glu Tyr Gly Arg Ala
210          215          220
Ala Asp Met Ile Glu Lys Arg Ile Lys Asp Phe Ile Glu Gly Ser Phe
225          230          235          240
Ser Asp Gly Val Thr Val Ser Glu Glu Asp Gly Gln Ser Tyr Val Val
245          250          255
Pro Thr Ile Ser Thr Val Leu Thr Asn Met Val Ser Val Ile Gln Glu

```

			260					265					270			
Gly	Phe	Tyr	Pro	Pro	Lys	Val	Gly	Ser	Phe	His	Glu	Ala	Leu	Leu	Gly	
		275					280					285				
Arg	Glu	Ile	Met	Val	Leu	Leu	Ser	Ala	Ala	Ile	Asp	Ala	Glu	Tyr	Arg	
	290					295					300					
Ala	Val	Leu	Ser	Arg	Thr	Arg	Asn	Ala	Lys	Pro	Asn	Pro	Leu	Thr	Thr	
305					310					315					320	
Lys	Leu	Asp	Lys	Tyr	Val	Asn	Asn	Pro	His	Leu	Gln	Met	Pro	Ser	Glu	
				325					330					335		
Ser	Val	Thr	Glu	Arg	Glu	Lys	Glu	Trp	Val	Glu	Arg	Glu	Arg	Glu	Arg	
			340					345					350			
Ile	Lys	Thr	Thr	Asp	Met	Thr	Ala	Glu	Asn	Leu	Phe	Arg	Asp	His	Pro	
		355					360					365				
Tyr	Leu	Pro	Lys	Ala	Ile	Asp	Gly	Ile	Leu	Gly	Pro	Lys	Arg	Thr	Pro	
	370					375					380					
Thr	Ala	Leu	Gln	Ala	Leu	Gln	Arg	Glu	Tyr	Lys	Arg	Cys	Asn	Lys	Phe	
385					390					395					400	
Asn	Asp	Ile	Val	Ser	Pro	Glu	Thr	Leu	Glu	Tyr	Phe	Leu	Val	Asn	Asn	
				405					410					415		
Arg	Gln	Val	Met	Phe	Ser	Asn	Tyr	Ser	Val	Thr	Arg	Val	Leu	Asp	Pro	
			420					425					430			
Asp	Ser	Ala	Ala	Arg	Phe	Ser	Met	Tyr	Val	Leu	Trp	Asn	Ala	Leu	Phe	
		435					440					445				
Leu	Cys	Ser	Gly	Gly	Leu	Thr	Gln	Lys	Thr	Asn	Ser	Ser	Ala	Val	Lys	
	450					455					460					
Ser	Arg	Leu	Ile	Leu	Gln	Val	Phe	Leu	Lys	Asp	Met	His	Ser	Leu	Phe	
465					470					475					480	
Val	Cys	Gln	Arg	Cys	Glu	Ser	Gly	Phe	Ile	Thr	Lys	Ser	Leu	Asp	Thr	
				485					490					495		
Phe	Thr	Ile	Ser	Leu	Lys	Glu	Gln	Ser	Lys	Pro	Ser	Met	Gly	Glu	Gln	
			500					505					510			
Glu	Leu	Glu	Thr	Tyr	Trp	Lys	Ala	Val	Leu	Asp	Ala	Leu	Gly	Gly	Gly	
		515					520					525				
Gly	Gly	Asn	Asn	Lys	Gly	Ala	Glu	Asn	Val	Asn	Gly	Leu	Gly	Glu	Leu	
	530					535					540					
Met	Val	Glu	Ile	Leu	Ser	Ala	Asp	Ser	Gly	Leu	Leu	Arg	Gly	Gly	Gly	
545					550					555					560	
Leu	Gly	Gly	Asp	Ile	Gly	Phe	Glu	Gly	Lys	Met	Lys	Gln	Lys	Arg	Glu	
			565						570					575		
Asp	Glu	Glu	Val	Arg	Asn	Met	His	Leu	Val	Asp	Lys	Lys	Gly	Tyr	Val	
			580					585					590			
Phe	Glu	Ala	Ala	Lys	Tyr	Val	His	Val	Ser	Lys	Gly	Phe	Ala	Ala	Leu	
		595					600					605				
Ser	Phe	Tyr	Leu	Leu	Tyr	Ala	Ala	Ala	Ala	Thr	Ser	Asn	Pro	Ser	Ile	
	610					615	</									

[illegible]

<210>	142
<211>	591
<212>	DNA
<213>	SHRIMP

<400> 142						
atggacggtt	cttcctataa	gagcactatt	gactaccaca	acattgaaga	tatggacgat	60
ctccagcgcg	ccacctacaa	ggatcgatat	gagacggaat	tggtcctcga	gatggctaag	120
aaggagggaa	ggtacgtccg	atcgtttggcc	accattggacg	aattggagggt	acctgaagaa	180
cgacagcatt	gctacacttg	cggctacacc	tttattagac	gcagggcacc	cccaccaaaa	240
ccgaagtcaa	tattcagaga	gccttgcgct	taccagaaac	ttctccccga	tgaccatcc	300
cccgtccgtt	tagaagagct	tgtcgacgtg	ccagaaggag	cgagtttttt	cacctacctt	360
ccctacgacg	acggatcttc	tacatcgtct	tcacaagccg	aatgtgaaga	tgattatcct	420
ccaccataat	accattcaga	aaatccacac	aggtcccaag	tgtgtgatta	ttgtaccaca	480
cgtaacagtc	tcagttctat	gacggatcac	gccagggccca	acgtctataaa	aaatctgaag	540
aqggaqaaqa	aqgccctqqq	tcttggccct	cqaacaaact	ttagctacta	q	591

<210>	143
<211>	194
<212>	PRT
<213>	SHRIMP

<400>	143														
Met	Asp	Val	Ser	Ser	Tyr	Lys	Ser	Thr	Ile	Asp	Tyr	His	Asn	Ile	Glu
1				5					10					15	
Asp	Met	Asp	Asp	Leu	Gln	Arg	Ala	Thr	Tyr	Lys	Asp	Arg	Met	Glu	Leu
			20					25					30		
Val	Leu	Glu	Met	Ala	Lys	Lys	Glu	Gly	Arg	Tyr	Val	Arg	Ser	Leu	Ala
		35					40					45			
Thr	Met	Asp	Glu	Leu	Glu	Val	Pro	Glu	Glu	Pro	Ala	Thr	Cys	Tyr	Thr
	50					55					60				
Cys	Gly	Tyr	Thr	Phe	Ile	Arg	Arg	Arg	Ala	Pro	Pro	Pro	Lys	Arg	Lys
65				70						75				80	
Ser	Ile	Phe	Arg	Glu	Pro	Cys	Ala	Tyr	Pro	Glu	Leu	Leu	Pro	Asp	Ala
				85					90					95	
Pro	Ser	Pro	Val	Arg	Leu	Glu	Glu	Leu	Val	Asp	Val	Pro	Glu	Gly	Ala
			100					105					110		
Ser	Phe	Phe	Thr	Tyr	Pro	Pro	Tyr	Asp	Asp	Gly	Ser	Ser	Thr	Ser	Ser
		115					120					125			
Ser	Gln	Ala	Glu	Cys	Glu	Asp	Asp	Tyr	Pro	Pro	Pro	Tyr	Asp	Pro	Ser
	130					135					140				
Glu	Asn	Pro	Gln	Arg	Ser	Gln	Val	Cys	Asp	Tyr	Cys	Thr	Thr	Arg	Gln
145					150					155				160	
Val	Leu	Ser	Ser	Met	Thr	Asp	His	Ala	Arg	Ala	Asn	Leu	Ile	Lys	Asn
				165					170					175	

Leu Lys Arg Glu Lys Lys Ala Leu Gly Leu Gly Arg Arg Asn Asn Phe  
 180 185 190  
 Ser Tyr

<210> 144  
 <211> 627  
 <212> DNA  
 <213> SHRIMP

<400> 144  
 atggcccca actccttcca gaaatttgct ccagttatca agactgagaa gaaggaagaa 60  
 gaaagggatg aacatgacga ccctttacgg cagattgatt ttagagatag aaagacatta 120  
 atctgcctca ctgcaaaactg tgtttcgagg aagagaaaaag ctggatctgc acatgatcga 180  
 gtatacaaaag tactacgcta cggaaccca tacaagtacc gtcgcccacaa tagaacacat 240  
 cgaggattgg ccctctcaat ggatcaagggt gaagtaggaa catgcctccc tctgcgaccc 300  
 atggaagaga ctgaagaaaa ccccatcgac aagtgcggag tggcgttcct gtactccaac 360  
 tacaatgaag gcgatggcat gaccacctt tacaacgacg aagagtatat aaagaagtgc 420  
 aaaacaattg aaggaggaac aagaacgtgg gtaaagaaga accgccaaga atacttcaga 480  
 caagctctag agacattgat gatgtcccat tctataaaac aatattccaa ttttattttt 540  
 ttcaaggagg atatggagga aggatttggtg cacaactcc acacatttat taatatggta 600  
 caccctaaaa aggtgtctgt tttgtga 627

<210> 145  
 <211> 204  
 <212> PRT  
 <213> SHRIMP

<400> 145  
 Met Ala Pro Asn Ser Phe Gln Lys Phe Ala Pro Val Ile Lys Thr Glu  
 1 5 10 15  
 Lys Lys Glu Glu Glu Arg Asp Glu His Asp Asp Pro Leu Arg Gln Ile  
 20 25 30  
 Asp Phe Arg Asp Arg Lys Thr Leu Ile Cys Leu Thr Ala Asn Cys Val  
 35 40 45  
 Ser Arg Lys Arg Lys Ala Gly Ser Ala His Asp Arg Val Tyr Lys Val  
 50 55 60  
 Leu Arg Tyr Gly Asn Pro Tyr Lys Tyr Arg Arg Pro Asn Arg Thr Gly  
 65 70 75 80  
 Leu Ala Leu Ser Met Asp Gln Gly Glu Val Gly Thr Cys Leu Pro Leu  
 85 90 95  
 Arg Pro Met Glu Glu Thr Glu Glu Asn Pro Ile Asp Lys Cys Gly Val  
 100 105 110  
 Ala Phe Lys Asn Tyr Asn Glu Gly Asp Gly Met Thr His Leu Tyr Asn  
 115 120 125  
 Asp Glu Glu Tyr Ile Lys Lys Cys Lys Thr Ile Glu Gly Gly Thr Arg  
 130 135 140  
 Thr Trp Val Lys Lys Asn Arg Gln Glu Tyr Phe Arg Gln Ala Leu Glu  
 145 150 155 160  
 Thr Leu Met Met Ser His Ser Ile Lys Gln Tyr Ser Asn Phe Ile Phe  
 165 170 175  
 Phe Lys Glu Asp Met Glu Glu Gly Phe Val His Lys Leu His Thr Phe  
 180 185 190  
 Ile Asn Met Val His Pro Lys Lys Val Ser Val Leu  
 195 200

<210> 146  
 <211> 870  
 <212> DNA

Met 1	Glu	Gly	Glu	His 5	Gln	Tyr	Leu	Asn	Leu 10	Val	Arg	Glu	Ile	Arg 15	Gly
Val	Lys	Lys	Asp 20	Asp	Arg	Thr	Gly	Thr 25	Gly	Thr	Leu	Ser	Ile 30	Phe	Gly
Pro	Gln	Met 35	Arg	Phe	Ser	Leu	Arg 40	Asp	Asp	Thr	Ile	Pro 45	Val	Leu	Thr
Thr	Lys 50	Lys	Ile	Phe	Trp 55	Arg	Gly	Val	Val	Glu	Glu 60	Leu	Leu	Trp	Phe
Ile 65	Arg	Gly	Asn	Thr 70	Asp	Ala	Lys	Glu	Leu	Ala 75	Lys	Lys	Lys	Ile	His 80
Ile	Trp	Asn	Ala	Asn 85	Gly	Ser	Arg	Glu	Phe 90	Leu	Asp	Ser	Arg	Gly 95	Asp
Arg	Ala	Glu	Gly 100	Asp	Leu	Gly	Pro	Val 105	Tyr	Gly	Phe	Gln	Trp 110	Arg	His
Phe	Gly	Ala 115	Glu	Tyr	Asp	Thr	Cys 120	Ser	Ser	Asp	Tyr	Thr 125	Gly	Lys	Gly
Ile	Asp 130	Gln	Leu	Ala	Asn 135	Ile	Leu	Lys	Thr	Leu	Arg 140	Glu	Asn	Pro	Asp
Asp 145	Arg	Arg	Met	Ile 150	Met	Thr	Ala	Trp	Asn	Pro 155	Met	Asp	Leu	His	Leu 160
Met	Ala	Leu	Pro	Pro 165	Cys	His	Met	Thr	Ala 170	Gln	Phe	Tyr	Val 175	Ala	Asn
Gly	Glu	Leu	Ser 180	Cys	Gln	Leu	Tyr	Gln 185	Arg	Ser	Gly	Asp	Val 190	Gly	Leu
Gly	Val	Pro 195	Phe	Asn	Ile	Ala	Ser 200	Tyr	Ser	Leu	Leu	Thr 205	His	Leu	Met
Ala	Ser 210	Met	Val	Gly	Leu	Lys 215	Pro	Gly	Glu	Phe	Ile 220	Leu	Thr	Leu	Gly
Asp 225	Ala	His	Ile	Tyr 230	Asn	Thr	His	Ile	Glu	Val 235	Leu	Lys	Lys	Gln	Leu 240
Cys	Arg	Val	Pro	Arg 245	Pro	Phe	Pro	Lys	Leu 250	Arg	Ile	Leu	Met	Ala 255	Pro
Glu	Lys	Ile	Glu 260	Asp	Phe	Thr	Ile	Asp 265	Met	Phe	Tyr	Leu	Glu 270	Gly	Tyr
Gln	Pro	His 275	Ser	Gly	Asn	Leu	Gln 280	Met	Lys	Met	Ala	Val 285			



WO 01/38351

229

PCT/US00/28888

<210> 148  
<211> 678  
<212> DNA  
<213> SHRIMP

<400> 148  
atggccttta attttgaaga ctctacaaat ctcttttgcca atatggactt gacggctggc 60  
acaacaacag accctacccg ccccaatata atattctttg aaagtctact ccccaactct 120  
ggtattgagg tgatgaagag gcgtctcgta cggcaaggaa agtgtgggaa ttttgaagca 180  
agtggaggtg ctatgtcgta tttctggctc gaagataatg cagaagatat ggagaatctc 240  
aacagtgggt cccatgtcaa gacaaactgc ttggcattat tccttcaaga gtttatcagc 300  
aactggattg aagagactga tgcacatgga cagtactgta cttttcccca atacatggac 360  
ggtggggatg gttcacgtgg gggatatttt acttcgctag ccatgaaatg gatggctagg 420  
gatgtgactt tctttgtgtt tgttgatagg aataatactg tagaaaatgc ggcattccata 480  
tggtatgtacc aaaaactact agcaatttgt gcaaaggtag taaaggatgat tgttgacaat 540  
gcatcaaacc caatgttttc tgtatgtaat gcgtgtagggt gcaagtaccc aggcccagtg 600  
tcatacgtaa ttgaaggcca tggagtgggt cattctgatt tgacatgtga tgagatttct 660  
ggattctttg tataataa 678

<210> 149  
<211> 220  
<212> PRT  
<213> SHRIMP

<400> 149  
Met Ala Phe Asn Phe Glu Asp Ser Thr Asn Leu Phe Ala Asn Met Asp  
1 5 10 15  
Leu Thr Ala Gly Thr Thr Thr Asp Pro Trp Asn Ile Ile Phe Phe Glu  
20 25 30  
Ser Leu Leu Pro Asn Ser Gly Ile Glu Val Met Lys Arg Arg Leu Val  
35 40 45  
Arg Gln Gly Lys Cys Gly Asn Phe Glu Ala Ser Gly Gly Ala Met Ser  
50 55 60  
Tyr Phe Trp Leu Glu Asp Asn Ala Glu Asp Met Glu Asn Leu Asn Ser  
65 70 75 80  
Gly Ser His Val Lys Thr Asn Cys Leu Ala Leu Phe Leu Gln Glu Phe  
85 90 95  
Ile Ser Asn Trp Ile Glu Glu Thr Asp Arg His Gly Gln Tyr Cys Thr  
100 105 110  
Phe Pro Gln Tyr Met Asp Gly Gly Asp Gly Ser Arg Gly Gly Tyr Phe  
115 120 125  
Thr Ser Leu Ala Met Lys Trp Met Ala Arg Asp Val Thr Phe Phe Val  
130 135 140  
Phe Val Asp Arg Asn Asn Thr Val Glu Asn Ala Ala Ser Ile Trp Met  
145 150 155 160  
Tyr Gln Lys Leu Leu Ala Ile Gly Ala Lys Val Val Lys Val Ile Val  
165 170 175  
Asp Asn Asn Pro Met Phe Ser Val Cys Asn Ala Cys Arg Cys Lys Tyr  
180 185 190  
Pro Gly Pro Val Ser Tyr Val Ile Glu Gly His Gly Val Gly His Ser  
195 200 205  
Asp Leu Thr Cys Asp Glu Ile Ser Gly Phe Phe Val  
210 215 220

<210> 150  
<211> 1269  
<212> DNA  
<213> SHRIMP

&lt;400&gt; 150

```

atggattcta atacttctat tttaccgcca agcaaacggc caggggttaaa tctgttacag 60
gttttaggga ttataataac ggtagcatta atagcttccg ttatcatcctt tatattttat 120
agggtaggta aacgcaaata ttacccttct tcactcctct cctcagaatt atctgatgta 180
gataatgggg tagaaggagg aggaggaaca acaacgacac caactcaacc ttcacctgac 240
ggtggagatg gatacgtaga tctttctcct caaaagaagg ctgaactaag aactagagtt 300
gcaaacgtca tctttcaaga agtgtcaaag gatcaaggag tggccttttag acgggcaatg 360
aatgattcaa ctgataagat aatggaagaa actgaggcga gaatcaataa cttttcagag 420
ccattcagag aagcaaccgt agaacgtgaa gtgtttaagg atgacacaga caaaaacttt 480
atcctttcaa ctctagattt aacagaggaa caattttaagg acattgttat ggctgaagtg 540
aaaaatcaat tagaaaattt tgactatgaa gacatgaccc gtctcatctt tgataacatc 600
ccagagactg attattttat gacaactcat ttcgatccga aaaaatatga cacgtactct 660
gaaaagggtat taggggttctc agatataaat agtatagaaa gaatatcctc tacattttat 720
aaaggtaaaa aatatgaggt aactactgga aatgtagctg tcctcgttga ttttgaatct 780
gaaacaataa aagagaaggc aggaaatagt ctcatccgta atgtcagatt tattgttggtg 840
gacgaacaga cctacaaatc tttcttccct gcattcaatc aagttttctt ctccttttaa 900
gtaaataagg agaaaaggga agttactgta tccatcaata atggatgtgt aggtatagtg 960
gccaatatta ctctctaac tacgccagtt ggagcagctt ccggacacta catctatggc 1020
actagcacag caaaggaaaa gacctatcta tttgtaatag acaagtacga taccactgaa 1080
tttgtttgtg gtctgagtaa caagtcaact cctctcatgg ctctaaatat tctctttatg 1140
agtgatactg ttttcccttc atttgacgaa gcagaaagac ctctgacgga tgccaaggca 1200
gtagaaattt taggtaaaag actaggtgta ggaagataca caaacgcaa catcagaaat 1260
actcagtga                                     1269

```

&lt;210&gt; 151

&lt;211&gt; 418

&lt;212&gt; PRT

&lt;213&gt; SHRIMP

&lt;400&gt; 151

```

Met Asp Ser Asn Thr Ser Ile Leu Pro Pro Ser Lys Arg Pro Gln Leu
1          5          10          15
Leu Gln Val Leu Gly Ile Ile Ile Thr Val Ile Ala Ser Val Ser Ser
20          25          30
Phe Ile Phe Tyr Arg Val Gly Lys Arg Lys Tyr Tyr Pro Ser Ser Ser
35          40          45
Ser Ser Ser Glu Leu Ser Asp Val Asp Asn Gly Val Glu Gly Gly Gly
50          55          60
Gly Thr Thr Thr Thr Pro Thr Gln Pro Ser Pro Asp Gly Gly Asp Gly
65          70          75          80
Tyr Val Asp Leu Ser Pro Gln Lys Lys Ala Glu Leu Arg Thr Arg Val
85          90          95
Ala Asn Val Ile Phe Gln Glu Val Ser Lys Asp Gln Gly Val Ala Phe
100         105         110
Arg Arg Ala Met Asn Asp Ser Thr Asp Lys Ile Met Glu Thr Glu
115         120         125
Ala Arg Ile Asn Asn Phe Ser Glu Pro Phe Arg Glu Ala Thr Val Glu
130         135         140
Arg Glu Val Phe Lys Asp Asp Thr Asp Lys Asn Phe Ile Leu Ser Thr
145         150         155         160
Leu Asp Leu Thr Glu Gln Phe Lys Asp Ile Val Met Ala Glu Val
165         170         175
Lys Asn Gln Leu Glu Asn Phe Asp Tyr Glu Asp Met Thr Arg Leu Ile
180         185         190
Phe Asp Asn Ile Pro Glu Thr Asp Tyr Leu Trp Thr Thr His Phe Asp
195         200         205
Pro Lys Lys Tyr Asp Thr Tyr Ser Glu Lys Val Leu Gly Phe Ser Asp
210         215         220
Ile Asn Ser Ile Glu Arg Ile Ser Ser Thr Phe Tyr Lys Gly Lys Lys
225         230         235         240

```

WO 01/38351

231

PCT/US00/28888

Tyr Glu Val Thr Thr Gly Asn Val Ala Val Leu Val Asp Phe Glu Ser  
 245 250 255  
 Glu Thr Ile Lys Glu Lys Ala Gly Asn Ser Leu Ile Arg Asn Val Glu  
 260 265 270  
 Phe Ile Val Val Asp Glu Gln Thr Tyr Lys Ser Phe Phe Pro Ala Phe  
 275 280 285  
 Asn Gln Val Phe Phe Ser Phe Lys Val Asn Lys Glu Lys Arg Glu Val  
 290 295 300  
 Thr Val Ser Ile Asn Asn Gly Cys Val Gly Ile Val Ala Asn Ile Thr  
 305 310 315 320  
 Pro Leu Thr Thr Pro Val Gly Ala Ala Ser Gly His Tyr Ile Tyr Gly  
 325 330 335  
 Thr Ser Thr Ala Lys Glu Lys Thr Tyr Leu Phe Val Ile Asp Lys Tyr  
 340 345 350  
 Asp Thr Thr Glu Phe Val Cys Gly Leu Ser Asn Lys Ser Thr Pro Leu  
 355 360 365  
 Met Ala Leu Asn Ile Leu Phe Met Ser Asp Thr Val Phe Pro Ser Phe  
 370 375 380  
 Asp Glu Ala Glu Arg Pro Leu Thr Asp Ala Lys Ala Val Glu Ile Leu  
 385 390 395 400  
 Gly Lys Arg Leu Gly Val Gly Arg Tyr Thr Asn Ala Asn Ile Arg Asn  
 405 410 415  
 Thr Gln

<210> 152  
 <211> 831  
 <212> DNA  
 <213> SHRIMP

<400> 152  
 atggaagggg ttatttttgga taagatagaa acaattgcaa aaagggcctc tcctagttat 60  
 ggggtctattg atgtgggtac ggctattttg cgccgtcaat tcatggaaaa aattaggggt 120  
 aaaataaatg aagaaaccac aatggagaag attatgggca caaaggaaga aagagaggac 180  
 actataagaa gtatagtggc taatgtttatc aaagagaata ctgttaaaga aaatgtaacc 240  
 gaaaaaatta gagcaatgac agataaggaa ttaaatgaca ataggggaatt tatgcatgat 300  
 tttggaaaaa tttcaactgg agatggagga accttccatc tctttgaaga tacaccgggt 360  
 tttgaaagtg ctttaaaggc agaataataa aacgttccag gagcaactac tccaaaatac 420  
 gtatctatga acagtttacg tatcgatgcg attaatggaa aaatcgaaga ggtttataat 480  
 ccttcaccta tcatgggtat tagagaatac ggcaccattc gcagggggcag gtacgaagaa 540  
 aatgcagggtt cgaaagaatt ggtttttatg accaagattg aaaaaagacc caataatgta 600  
 gctgaaaatc tcattattag agttgccaac cagcagtata atgttatgag gatgggtgtt 660  
 tttatagact acgaaacaaa gaaggggggtg tccaaggagg aaatgtttat accatataat 720  
 gttcagaaaa caaaggctct taaagggcgt agtacttact tttcattcgt aaggaaaatt 780  
 cctgatgaac cagaaggagg tatcataata cacgcactag ggtttttattg a 831

<210> 153  
 <211> 274  
 <212> PRT  
 <213> SHRIMP

<400> 153  
 Met Glu Gly Val Ile Leu Asp Lys Ile Glu Thr Ile Ala Lys Arg Asp  
 1 5 10 15  
 Ser Tyr Gly Ser Ile Asp Val Gly Thr Ala Ile Leu Arg Arg Gln Phe  
 20 25 30  
 Met Glu Lys Ile Arg Gly Lys Ile Asn Glu Glu Thr Thr Met Glu Lys  
 35 40 45  
 Ile Met Gly Thr Lys Glu Glu Arg Glu Asp Thr Ile Arg Ser Ile Val  
 50 55 60

[illegible]

<210>	154
<211>	486
<212>	DNA
<213>	SHRIMP

<400> 154						
atgacatctc	cagctccatc	accctcttcc	acccccaaat	ccagttgtac	cactattgta	60
aaccgatgtg	gtttcctcct	tgacaacaac	aagggaagtgg	tcatctacga	caccaattcc	120
aaattcaagt	gtgaacccaa	aaatctggaa	ctaattgggtg	tactttctgg	agtctctgat	180
aattgttgta	ccgagatata	ccccgaccag	atatttgttg	gaacatatat	ggtcaaatat	240
aactgggtcta	aatctgggtca	tgaacgcttc	agtgcacatga	gtaacacaatg	tctggacaat	300
attacacgcc	cttcagaagt	gattgaaagt	gtgataaaga	aaacgtccag	cgactttaaa	360
atgaagtaca	cacgttcctt	gatggaccac	accgagaaat	actatTTTTc	tgytgaccaa	420
aaattgagca	aaattagtag	ttggtgtaca	accctataac	gacagtggggt	atgcaactcc	480
gcttag						488

<210>	155
<211>	159
<212>	PRT
<213>	SHRIMP

<400> 155															
Met	Thr	Ser	Pro	Ala	Pro	Ser	Pro	Ser	Ser	Thr	Pro	Lys	Ser	Ser	Cys
1				5					10					15	
Thr	Thr	Ile	Val	Asn	Arg	Cys	Gly	Phe	Leu	Leu	Asp	Asn	Asn	Lys	Glu
			20					25					30		
Val	Val	Ile	Tyr	Asp	Thr	Asn	Ser	Lys	Phe	Lys	Cys	Glu	Pro	Lys	Asn
		35					40					45			
Leu	Glu	Leu	Ile	Gly	Val	Leu	Ser	Gly	Val	Ser	Asp	Asn	Val	Val	Thr
	50					55					60				
Gln	Ile	Ser	Pro	Asp	Gln	Ile	Phe	Val	Gly	Thr	Tyr	Met	Val	Lys	Tyr

WO 01/38351

233

PCT/US00/28888

65	Asn	Trp	Ser	Lys	Ser	Gly	His	Glu	Arg	Phe	Ser	Asp	Met	Ser	Asn	Asn
				85						90					95	
Cys	Leu	Asp	Asn	Ile	Trp	Ser	Glu	Val	Ile	Glu	Ser	Val	Ile	Lys	Lys	
			100					105						110		
Thr	Ser	Ser	Asp	Phe	Lys	Met	Lys	Tyr	Thr	Arg	Ser	Leu	Met	Asp	His	
			115				120					125				
Thr	Glu	Lys	Tyr	Tyr	Phe	Ser	Gly	Asp	Gln	Lys	Leu	Ser	Lys	Ile	Ser	
			130				135				140					
Ser	Trp	Cys	Thr	Thr	Pro	Ile	Arg	Gln	Trp	Val	Cys	Asn	Ser	Val		
145					150					155						

<210> 156  
 <211> 894  
 <212> DNA  
 <213> SHRIMP

<400> 156

atggcattac	aggaaaagga	tataactata	gggaatgttt	ctgctgccct	acgagagttg	60
atgtactcac	ccacacatat	gcagcatcac	gataagctaa	acacattcct	ggacagaaat	120
gttgaatcat	cttcagaaga	gaaaataaga	caaattgttg	ataaaatacg	atcccaaaca	180
acatctgaca	tatctgaaac	agtcaataat	gtcacaaacta	atgggactgc	attttccctt	240
ttcgaagata	ccttagaagg	tatggtgaaa	aaaaatatag	gtgataacct	tcagagtggg	300
gacttttattg	atggccgtaa	aaagctcaat	gacatgaaga	gtctagctac	tggagccatc	360
ttatctagac	agcgagattt	tgttgcagaa	agtataacag	gaacaaagga	ctgggtcaag	420
gctataatgg	gttgtggtat	tataagggtat	actgtatttg	tcaataacct	tgcaagatca	480
acactcgata	atgatgatga	caaggcagca	acctattata	acacccttat	atatggcggg	540
tattgtaaaa	tggtataaaa	ggactatgaa	ataccagatt	cgtacagcaa	ggtcgaagcg	600
gaacatacac	ttgaagggaag	aaagatgacc	tttaatatata	aatggagagg	cgataccata	660
aacaacctaa	taacaatcat	cccttcagtg	acaggttatc	ttgcttccat	ctctgaagac	720
gcagatgtgc	aggcgccatt	acttttaaac	tgcaacaact	gttttataga	ggcagatatg	780
agtagcctct	acatggatga	gaaaaaaaca	gaggcatcat	ttaccctcaa	cttaccggaa	840
atcgaaggag	ctgatgcgaa	tgcagtcctat	gaaatatgta	tagtagtagt	ttga	894

<210> 157  
 <211> 297  
 <212> PRT  
 <213> SHRIMP

<400> 157

Met	Ala	Leu	Gln	Glu	Lys	Asp	Ile	Thr	Ile	Gly	Asn	Val	Ser	Ala	Ala
1				5					10					15	
Leu	Arg	Glu	Leu	Met	Tyr	Ser	Pro	Thr	His	Met	Gln	His	His	Asp	Lys
			20					25					30		
Leu	Asn	Thr	Phe	Leu	Asp	Arg	Asn	Val	Glu	Ser	Ser	Ser	Glu	Glu	Lys
			35				40					45			
Ile	Arg	Gln	Ile	Val	Asp	Lys	Ile	Arg	Ser	Gln	Thr	Thr	Ser	Asp	Ile
			50			55					60				
Ser	Glu	Thr	Val	Asn	Asn	Val	Thr	Thr	Asn	Gly	Thr	Ala	Phe	Ser	Leu
65						70				75				80	
Phe	Glu	Asp	Thr	Leu	Glu	Gly	Met	Val	Lys	Lys	Asn	Ile	Gly	Asp	Asn
				85					90					95	
Leu	Gln	Ser	Gly	Asp	Phe	Ile	Asp	Gly	Arg	Lys	Lys	Leu	Asn	Asp	Met
			100					105					110		
Lys	Ser	Leu	Ala	Thr	Gly	Ala	Ile	Leu	Ser	Arg	Gln	Arg	Asp	Phe	Val
			115				120					125			
Ala	Glu	Ser	Ile	Thr	Gly	Thr	Lys	Asp	Trp	Leu	Lys	Ala	Ile	Met	Gly
			130				135				140				
Cys	Gly	Ile	Ile	Arg	Tyr	Thr	Val	Phe	Val	Asn	Asn	Leu	Ala	Arg	Ser
145					150					155					160

Thr Leu Asp Asn Asp Asp Asp Lys Ala Ala Thr Tyr Tyr Asn Thr Pro  
 165 170 175  
 Ile Tyr Gly Gly Tyr Cys Lys Met Ala Ile Lys Asp Tyr Glu Ile Pro  
 180 185 190  
 Asp Ser Tyr Ser Lys Val Glu Ala Glu His Thr Val Glu Gly Arg Lys  
 195 200 205  
 Met Thr Phe Asn Ile Lys Trp Arg Gly Asp Thr Ile Asn Asn Leu Ile  
 210 215 220  
 Thr Ile Ile Pro Ser Val Thr Gly Tyr Leu Ala Ser Ile Ser Glu Asp  
 225 230 235 240  
 Ala Asp Val Gln Ala Pro Leu Leu Leu Asn Cys Asn Asn Cys Phe Ile  
 245 250 255  
 Glu Ala Asp Met Ser Ser Leu Tyr Met Asp Glu Lys Lys Thr Glu Ala  
 260 265 270  
 Ser Phe Thr Leu Asn Leu Pro Glu Ile Glu Gly Ala Asp Ala Asn Ala  
 275 280 285  
 Val Tyr Glu Ile Cys Ile Val Val Val  
 290 295

<210> 158  
 <211> 6588  
 <212> DNA  
 <213> SHRIMP

<400> 158  
 atgcaactcc gtctagaaaa ctttgtcaag gaagaacatg aaactgtcgt ggtgcacaac 60  
 ctttctggaa tgactggatt caacatattt aatagtctcc ccgtgtattt tgaagtgcac 120  
 aatgagatgg acgccctaatt ttttatggcg gctttcttga agcacaatag tttatgggga 180  
 gaaattaacg ccaatatgga cttgtacacg tttgattatg cgggtgcttt tctggacgaa 240  
 agatggtgcc accacgagaa gagtttttct gtcgtccgag cacaacttat caactcgtat 300  
 tacaagtgcg ggagaaaaat catgcaagcc ctggacaata actacaacaa caagaataag 360  
 aagaggaaga atgttggtgg agcacctgcg ttcacattta tgagcgggga cggagaggga 420  
 ggaaaggaag ccctagaagc tagtttcgat gtgattgggg gaacaagagg aggaagattt 480  
 ggtgttgatt caacaccatg ccccatcttct tcagccatgc aactaaaact ggacaatgaa 540  
 ggaaactatg gatgtattgc ctgcttcgca tcaatgttct ttgtattgga gaaccagggt 600  
 gatgaatctt cttcatatc aacggatgcc tctaaaattg gacaagcgca agcatggata 660  
 gatgaacgac tacgaaacaa tgaaaatgga ggagaagaaa ataatgtctt taaaaagacc 720  
 ttccatagtc tggctgatat taccctaaaag gctcatgaaa ctgcctattc caataccatc 780  
 ccacttgga ccaatggcag gcagtggaat tggcctactc acactgtgga acctattgcc 840  
 catgaatttg ttaccatttc tctagtaaac acattgaaaa atctagggga tagaaaactt 900  
 cccgatttca attttgatat cttgtacaac ttgtctaact cattttgaaa aatggtgcta 960  
 gtgtttattc aaaattgtca cattttaact ggacataaaa acaatgaaaa tgtggtgcct 1020  
 cgaggttctg cttctgggaa gtggtggact attaatgttg tgggtgtgaa catgtggact 1080  
 tttcaagtaa caaatgttaa agttgaaaag gatagaaaaa tatccgattt ggcctgtatg 1140  
 gaaactctcc ctgcgtctacc taatccagga agcactaccg tcgatgacag aatagttttt 1200  
 aagggattct gtagagggga aaatctaggg agtgtagggt aagtcgtatc cgacattaca 1260  
 cagagtgtca agaatttttg tctcatggtt gaaaatagga aatttagtgt ggataaagaa 1320  
 actggtttca tctcttcaga atcgatagtg tctgatccct tcttttctact agaagtgact 1380  
 ggctgtagat ctaatcgtgc ccaagatact attaataatg gccgagttag tgctcgtgta 1440  
 atgaggatcc taaagtcacg tgaaggtgct cgtgtatggt tggccaagga tgaaaatgcc 1500  
 atcatctttg aaaacgtaa ccacgatacg gccatctcta cggacgctat ggagcgagct 1560  
 ataggcgagc acaagatact gtactatgat attgaaacaa cagataaaga tttcaccgac 1620  
 aaaaaatcag tcatcacatc tattgggttc tgtttgtgta cgggaggcga tatgacacat 1680  
 ggaggagaga gaggagtatt tggactggtt gcacctggat ccgacgtgga aaaggtgaaa 1740  
 gagactataa taaattcgta cgatcctgaa gaaaaggaag acattatgaa acagtgcctt 1800  
 caagtgattg aaattttcac caacgagttt gaaatgttgc ttggttttgg aaagtatata 1860  
 gataaagtga agcctcacgt gattagtggg tggaacaatg tagcttttga cgaccctttt 1920  
 gtctttactc gtatcgtcaa acattttgag gtacacacca aagacatgtc ttattgtgta 1980  
 gcagatgcat ctacagcaga atctgtcctt cctagagcaa cagaaggagg aggaggagga 2040  
 gaaactccat atagattgag caccctcaa gaaagaatac aactagcaag cactggtatt 2100

tcaataaat	tgggaaaatt	tgtagacaag	aaaactggca	tgttgaacc	tgaatgact	2160
gcagatttat	tggccggggc	agaaagtcag	gccaatacca	agtttaagga	acgcaacaag	2220
ttatcctcca	gtaataaagg	atcagcagga	tggttccaga	aaattattgg	cggtagtgct	2280
agtgcctatc	ggttggtatc	catgaaagt	tgcgaaaagg	cctataaga	atccctctct	2340
gaatttaatt	tgaacgcgct	gctcgccaaa	tgtgagtgtg	tcggcgacaa	ggttaaaaat	2400
gtaaaagatg	aagtagacct	acactttcat	ctattgggat	tcttgaagct	gaagaaggcc	2460
caggatcagg	caaaagtaca	cgtctattgt	tgcaaggatg	cctacttgac	tggtatagtt	2520
tctacctcca	tcaacaagga	aggggagatt	tttaggctgt	gtatggactc	tgctttaacc	2580
gaggcggtcg	tgacagccaa	cctggccact	cctctatgta	taggaagaag	agcaatctgt	2640
agaaatatgg	gagaagaaag	ggcagataga	agaggtgtgg	taggtaagaag	acactctatt	2700
gccacagaca	caaagggagg	tatggtgagt	caacctatcg	tcaatcatgt	tccctatcaa	2760
acgattgaca	tgacaagttt	gtacccgatg	accatgtgtc	agaataatct	gtgcaccact	2820
accttttgtga	cccatcgaca	aattatgcaa	ctgagggata	gattgggtact	tgaaaaaatg	2880
aaaaacaaaa	ccaccgactc	ttttattgtt	tggagcttta	tgacgaagt	caatcagatt	2940
gtgttgtccg	agtcagagcc	cattgatatt	ctgagtcgat	catggaagaa	tagcaactct	3000
aatagacaaa	ctccaattac	tcgcatagag	gaaagtttgg	gtctaagatt	catagaaaat	3060
ttggatgccg	agaagacaaa	taataaaacg	tggtgcacca	atacatctcc	caatatgaat	3120
gtcactgccg	caggtatgga	ttacttcccc	gagatttgtt	gtgcacattaa	tatgcagttt	3180
gcggccaagg	tgaattatga	tatgcataata	gccccagcaa	gtttagagta	tatgcttcaa	3240
gtattgcccc	taatgttaat	cgacagacgg	tacattggcg	cacacataac	agctggaaaa	3300
tgtcgtacat	tggaggatat	tctctcagaa	ctcgaaaagg	acttttctgt	tgaaaaagat	3360
gaggaaatta	taagaaccca	ttggacattt	aagggtcaaa	aacaatacga	tttctgtcat	3420
agtctgtgta	cccaaatggc	tcgtcacatt	attgaatcta	cgggaagaaa	tatcctgtat	3480
tatgaaggca	atgaaaaatt	cgagaggttg	gttagcttat	cagacagaat	ttatcgccgc	3540
gttggcgcat	ttgattcggc	caatgatcca	gctgttagac	tgtgttcttc	tcgtctaatc	3600
aatgttgga	tgtttggttag	gacatggaac	gtaaaaactg	acatttcttaa	gggaatcatc	3660
cctcaaattgc	aagccactta	cagagccgat	cgagtttgtga	tgcaaacaa	ggccaaggag	3720
tttgccaaga	tgggagacat	gaaaagagct	ggtctaaaca	aagttggaca	aaatattatg	3780
aagctcggta	tgaattccat	gtatggtcat	ttggccctaa	gagcagcttc	gagccgtaaa	3840
ggaattggcg	ctggatctgc	caatactgcc	tcaagtattt	ccaacatgtc	agccacggga	3900
ggaattggag	gaggcacaag	gcactcgggtg	acggccaatc	agattacaga	aaatgctcga	3960
tgtgtatttg	gcaatatttg	ttgtggatta	cagatggctc	ttcctggtac	taagcagacg	4020
tacggggata	cagattctgt	attctgcgtg	cataatattg	tagtgtatgg	aggaatgata	4080
ccagaatatg	atgaacaaac	ttggcaaatat	tattatgtga	tggaattatgc	tctaaaaaat	4140
aaaattggctg	caattattcc	catctatgac	aactcgttaa	caaaggcact	ccagtttgtga	4200
gagcgccgag	acgctggtgt	gggcatgatg	aatatcgccc	atgaagctct	agctgtcgct	4260
ggtcttttgt	ttgccaagaa	aacataccat	atgcttcact	ttaatgaaaa	tagtgcagcg	4320
ttcaatgaca	tgataaaaatt	gaaatcaacc	gataacaata	ataagtttgc	atcctttatc	4380
aagagaccta	gccacgcaga	tgggtatggt	gttccccata	atccttcatt	gattctttaga	4440
gcggccgaag	gacccgctgg	taagaaaattg	aaaagctttt	tggaaagga	aggaattcat	4500
gcgagagaaga	gtatggagga	atggtttacc	tcttcacct	cgtgagtggc	catggtgct	4560
tctgttatca	acaacttgta	tgctcacaa	attgtagggg	tggaagagg	taactggatt	4620
gacgccatga	cttcccgccc	catagaagcg	ggtacagaaa	tgatgaggcc	ggtgacgcaa	4680
gcgaattgac	ctttcacccc	ttcaaaaaag	ggagcctttg	tgaagaaggg	aattacacc	4740
accaccaaac	taaaggtct	ccaatcatgt	attgcaagat	ttttaccaa	aatagaggaa	4800
aagaaaagct	gttatttgga	tgtgtgaag	aatcatgttg	agaattttgc	atcccatata	4860
acaaatcctg	ctatgatgat	tactagtctc	cgagtcaaca	agtttgatac	gtcaaaagaa	4920
cagagtagac	ctaactctct	agctctagcg	ataaataacc	acctgaacc	ttcttcagaa	4980
atttcatattg	ggcagaaatt	taagacgggtg	acatcagttt	cttcttgga	tctttcgcca	5040
gaggaagggg	aagtcctctg	tggttatttt	acagctggta	gcgtgcgttg	ggatgccacc	5100
aaactgaagg	gaagtgttcc	tgcattttca	gtcaagaatt	tatctgttgt	gccaacgcc	5160
atcacatctg	tatacaagat	ggttgagagc	gataagacgg	caataaaatc	catgattgct	5220
aaaaatgtag	aagtgtttgtg	ttctacatct	gccaatactg	gattttctct	gagaagagga	5280
gcattgtcat	ttaatacagg	cgctcattgtt	acaaaggacg	tggtctatggc	ttgtatacga	5340
tctctaaata	ataaacaaaat	gttattgttt	gttggagggg	gaaaggatta	cggtagaag	5400
gcgacgcagc	acgacgaaga	agcagaagaa	gaggacgaag	aaaattggtga	aaacgaagag	5460
aacaaaagggtg	actgtgtcac	ggaaaaaaaag	atccctggac	gaagcactaa	caaggatgtt	5520
ggtgaagaaa	ctaaaacaag	cgagaaaacg	gagggagaaa	gaaagggctc	taagacggca	5580
aaggggaaaga						

```

atgaaattgg accaaagagt agcagaccaa atggaaaatt taatatctca attggatcaa 5820
atacgtaatc tctccaacaa gaagaggcaa gaaaaggag ggccttttaa gtctgaattg 5880
gacgccatgg ttgctgcagt taaggttaag tttttcccag ttttagatgc gtctagaaaa 5940
ttgactcaag accattggaa aaagtgcgcc gtgtccatcc cagaaacgcg tgaagaaaaa 6000
ccattaatgg gtgtgccttt tgaagttgca ctcaattctc taataggaaa acacaagtgc 6060
acagatacat gcgacatggc ttgttgtcaa tcattgtatt ttgtcctctt gtacacgcta 6120
gctttaaaat ttgagaacga aagattggcc cggcaaattg gcctagatga ctctgtagat 6180
ttgatggctg agatgttggt cggaggggat aaactattgg ccaggaagt gttaaaaagg 6240
gtaaaagatg ctcaagatag aaagtgggtg aaatctttat tgcctttaaa ttataaccat 6300
gacacaaata caattatatt tttgtttgag tctttaaggt ttgtcagaa acctgtagct 6360
ggtagtgatg ttagtgaaat aaaagacgct gttagaggtc tggccttttc taccactaca 6420
ggtagtgatg ggaattatac tgatgaaaga ttttttgac cattgtataa catggatgaa 6480
ctttgtaacg aacgtgtcaa tggaaattgt aaattgtcct ttataactgg tatttatcat 6540
acggcagcag tagaattggc tgctgcatgt ctatcttggtg ttttgtaa 6588

```

<210> 159  
 <211> 2183  
 <212> PRT  
 <213> SHRIMP

<400> 159

Met	Gln	Leu	Arg	Leu	Glu	Asn	Phe	Val	Lys	Glu	Glu	His	Glu	Thr	Val
1				5					10				15		
Val	Val	His	Asn	Pro	Ser	Gly	Met	Thr	Gly	Phe	Asn	Ile	Phe	Asn	Ser
			20					25				30			
Ser	Pro	Val	Tyr	Phe	Glu	Val	His	Asn	Glu	Met	Asp	Ala	Leu	Ile	Phe
		35					40				45				
Met	Ala	Ala	Phe	Leu	Lys	His	Asn	Ser	Leu	Trp	Gly	Glu	Ile	Asn	Ala
	50					55					60				
Asn	Met	Asp	Leu	Tyr	Thr	Phe	Asp	Tyr	Ala	Gly	Ala	Phe	Leu	Asp	Glu
65					70					75				80	
Arg	Trp	Cys	His	His	Glu	Lys	Ser	Phe	Ser	Val	Val	Arg	Ala	Gln	Leu
			85					90						95	
Ile	Asn	Ser	Tyr	Tyr	Lys	Cys	Arg	Arg	Lys	Ile	Met	Gln	Ala	Leu	Asp
			100					105					110		
Asn	Asn	Tyr	Asn	Asn	Lys	Asn	Lys	Lys	Arg	Lys	Asn	Val	Gly	Gly	Ala
		115					120					125			
Pro	Ala	Phe	Thr	Phe	Met	Ser	Gly	Asp	Gly	Glu	Gly	Gly	Lys	Glu	Ala
	130					135					140				
Leu	Glu	Ala	Ser	Phe	Asp	Val	Ile	Gly	Gly	Thr	Arg	Gly	Gly	Arg	Phe
145					150					155				160	
Gly	Val	Asp	Ser	Thr	Pro	Cys	Pro	His	Ser	Ser	Ala	Met	Gln	Leu	Lys
			165					170						175	
Leu	Asp	Asn	Glu	Gly	Asn	Tyr	Gly	Cys	Ile	Ala	Cys	Phe	Ala	Ser	Met
		180					185					190			
Phe	Phe	Val	Leu	Glu	Asn	Pro	Gly	Asp	Glu	Ser	Ser	Phe	Ile	Ser	Thr
	195					200						205			
Asp	Ala	Ser	Lys	Ile	Gly	Gln	Ala	Gln	Ala	Trp	Ile	Asp	Glu	Arg	Leu
	210					215					220				
Arg	Asn	Asn	Glu	Asn	Gly	Gly	Glu	Glu	Asn	Asn	Val	Phe	Lys	Lys	Thr
225					230					235				240	
Phe	His	Met	Leu	Ala	Asp	Ile	Thr	Gln	Lys	Ala	His	Glu	Thr	Ala	Tyr
			245					250						255	
Ser	Asn	Thr	Ile	Pro	Leu	Gly	Pro	Asn	Gly	Arg	Gln	Trp	Asn	Trp	Pro
		260						265					270		
Thr	His	Thr	Val	Glu	Pro	Ile	Ala	His	Glu	Phe	Val	Thr	His	Ser	Leu
	275						280					285			
Val	Asn	Thr	Leu	Lys	Asn	Leu	Gly	Asp	Arg	Lys	Leu	Pro	Arg	Phe	Asn
	290					295					300				
Phe	Asp	Ile	Leu	Tyr	Asn	Leu	Leu	Asn	Pro	Phe	Gly	Lys	Met	Leu	Leu
305					310					315				320	



Val	Phe	Ile	Gln	Asn 325	Cys	His	Ile	Leu	Thr 330	Gly	His	Lys	Asn 335	Glu
Asn	Val	Val	Pro 340	Arg	Gly	Ser	Ala	Ser 345	Gly	Lys	Trp	Trp	Thr 350	Asn
Phe	Val	Gly 355	Val	Asn	Met	Trp	Thr 360	Phe	Gln	Val	Thr	Lys 365	Cys	Val
Glu	Lys	Asp 370	Arg	Lys	Ile	Ser 375	Asp	Leu	Ala	Cys	Met 380	Leu	Pro	Arg
Pro 385	Asn	Pro	Gly	Ser	Thr 390	Thr	Val	Asp	Asp	Arg 395	Ile	Val	Phe	Lys
Phe	Cys	Arg	Gly	Glu 405	Asn	Leu	Gly	Ser	Val 410	Gly	Glu	Val	Val	Ser
Ile	Thr	Gln	Ser 420	Val	Lys	Asn	Phe	Cys 425	Leu	Met	Val	Glu	Asn 430	Arg
Phe	Ser	Val 435	Asp	Lys	Glu	Thr	Gly 440	Phe	Ile	Ser	Ser	Glu 445	Ser	Ile
Ser	Asp	Pro	Phe	Phe	Ser	Leu 455	Glu	Val	Thr	Gly	Cys 460	Arg	Ser	Asn
Ala 465	Gln	Asp	Thr	Ile	Asn	Asn 470	Gly	Arg	Val	Ser	Ala	Arg	Val	Met
Ile	Leu	Lys	Ser	Arg 485	Glu	Gly	Ala	Arg	Val 490	Trp	Leu	Ala	Lys	Asp
Asn	Ala	Ile	Ile 500	Phe	Glu	Asn	Val	Asn 505	His	Asp	Thr	Ala	Ile 510	Ser
Asp	Ala	Met 515	Glu	Arg	Ala	Ile	Gly 520	Gln	His	Lys	Ile	Leu 525	Tyr	Tyr
Ile	Glu	Thr 530	Thr	Asp	Lys	Asp 535	Phe	Thr	Asp	Lys	Lys	Ser	Val	Ile
Ser 545	Ile	Gly	Phe	Cys 550	Leu	Cys	Thr	Gly	Gly	Asp 555	Met	Thr	His	Gly
Glu	Arg	Gly	Val	Phe 565	Gly	Leu	Val	Ala	Pro 570	Gly	Ser	Asp	Val	Glu
Val	Lys	Glu	Thr 580	Ile	Ile	Asn	Ser	Tyr 585	Asp	Pro	Glu	Glu	Lys 590	Glu
Ile	Met	Lys 595	Gln	Cys	Pro	Gln	Val 600	Ile	Glu	Ile	Phe	Thr 605	Asn	Glu
Glu 610	Met	Leu	Leu	Gly	Phe	Gly 615	Lys	Tyr	Ile	Asp	Lys	Val	Lys	Pro
Val 625	Ile	Ser	Gly	Trp	Asn 630	Asn	Val	Ala	Phe	Asp 635	Asp	Pro	Phe	Val
Thr	Arg	Ile	Val	Lys 645	His	Leu	Ser	Asp	His 650	Thr	Lys	Asp	Met	Ser
Cys	Val	Ala	Asp 660	Ala	Ser	Thr	Ala	Glu 665	Ser	Val	Leu	Pro	Arg 670	Ala
Glu	Gly	Gly 675	Gly	Gly	Gly	Glu	Thr 680	Pro	Tyr	Arg	Leu	Ser	Thr 685	Pro
Glu	Arg	Ile 690	Gln	Leu	Ala	Ser 695	Thr	Gly	Ile	Phe	Asn 700	Lys	Leu	Gly
Phe 705	Val	Asp	Lys	Lys	Thr 710	Gly	Met	Leu	Lys	Pro 715	Glu	Met	Thr	Ala
Leu	Leu	Ala	Gly	Ala 725	Glu	Ser	Gln	Ala	Asn 730	Thr	Lys	Phe	Lys	Glu
Asn	Lys	Leu	Ser 740	Ser	Ser	Asn	Lys	Gly 745	Ser	Ala	Gly	Trp	Phe 750	Gln
Ile	Ile	Gly 755	Gly	Met	Cys	Ser	Ala 760	Ile	Arg	Leu	Asp	Leu 765	Met	Lys
Cys	Glu	Lys 770	Ala	Tyr	Lys	Glu 775	Ser	Leu	Ser	Glu	Phe 780	Asn	Leu	Asn
Val 785	Leu	Ala	Lys	Val	Ser 790	Ser	Val	Gly	Asp	Lys 795	Val	Lys	Asn	Val
Asp	Glu	Val	Asp	Leu	His	Phe	His	Leu	Leu	Gly	Phe	Leu	Lys	Leu

					805					810					815	
Lys	Ala	Gln	Asp	Gln	Ala	Lys	Val	His	Val	Tyr	Cys	Cys	Lys	Asp	Ala	
			820					825					830			
Tyr	Leu	Thr	Gly	Ile	Val	Ser	Thr	Ser	Ile	Asn	Lys	Glu	Gly	Glu	Ile	
			835					840					845			
Phe	Arg	Leu	Cys	Met	Asp	Ser	Ala	Leu	Thr	Glu	Ala	Val	Val	Thr	Ala	
			850				855					860				
Asn	Leu	Ala	Thr	Pro	Leu	Cys	Ile	Gly	Glu	Gly	Ala	Ile	Cys	Arg	Asn	
865					870					875					880	
Met	Gly	Glu	Glu	Arg	Ala	Asp	Arg	Arg	Gly	Val	Gly	Val	Arg	Arg	His	
				885					890						895	
Ser	Ile	Ala	Thr	Asp	Thr	Lys	Gly	Gly	Met	Val	Ser	Gln	Pro	Ile	Val	
				900				905					910			
Asn	His	Val	Pro	Tyr	Gln	Thr	Ile	Asp	Met	Thr	Ser	Leu	Tyr	Pro	Met	
			915					920					925			
Thr	Met	Cys	Gln	Asn	Asn	Leu	Cys	Thr	Thr	Thr	Phe	Val	Thr	Gln	Ile	
						935						940				
Met	Gln	Leu	Arg	Asp	Arg	Leu	Val	Leu	Glu	Lys	Met	Lys	Asn	Lys	Thr	
945					950					955					960	
Thr	Asp	Ser	Leu	Leu	Leu	Leu	Asp	Val	Ile	Asp	Glu	Cys	Asn	Gln	Ile	
				965					970						975	
Val	Leu	Ser	Glu	Tyr	Arg	Pro	Ile	Asp	Ile	Ala	Val	Ala	Ser	Trp	Lys	
				980				985					990			
Asn	Ser	Asn	Ser	Asn	Arg	Gln	Thr	Pro	Ile	Thr	Arg	Ile	Glu	Glu	Ser	
							1000						1005			
Leu	Gly	Leu	Arg	Phe	Ile	Glu	Asn	Leu	Asp	Ala	Glu	Lys	Thr	Asn	Asn	
						1015					1020					
Lys	Thr	Trp	Cys	Thr	Asn	Thr	Ser	Pro	Asn	Met	Asn	Val	Thr	Ala	Ala	
1025					1030					1035					1040	
Gly	Met	Asp	Tyr	Phe	Pro	Glu	Ile	Val	Cys	Asp	Ile	Asn	Met	Gln	Phe	
				1045					1050						1055	
Ala	Ala	Lys	Val	Asn	Asp	Asp	Met	His	Ile	Ala	Pro	Ala	Ser	Leu	Glu	
				1060				1065						1070		
Tyr	Met	Leu	Gln	Val	Leu	Pro	Ile	Met	Leu	Ile	Asp	Arg	Pro	Tyr	Ile	
				1075				1080					1085			
Gly	Ala	His	Ile	Thr	Ala	Gly	Lys	Cys	Arg	Thr	Leu	Glu	Asp	Ile	Leu	
						1095					1100					
Ser	Glu	Leu	Glu	Lys	Asp	Phe	Ser	Val	Glu	Lys	Asp	Glu	Glu	Ile	Ile	
1105					1110					1115					1120	
Arg	Thr	His	Trp	Thr	Phe	Lys	Gly	Gln	Lys	Gln	Tyr	Asp	Phe	Cys	His	
				1125					1130						1135	
Ser	Pro	Val	Thr	Gln	Met	Ala	Arg	His	Ile	Ile	Glu	Ser	Thr	Gly	Arg	
				1140												

Gly Thr Arg His Ser Val Thr Ala Asn Gln Ile Thr Glu Asn Ala Arg  
 1300 1305 1310  
 Cys Val Phe Gly Asn Ile Gly Cys Gly Leu Gln Met Ala Leu Pro Gly  
 1315 1320 1325  
 Thr Lys Gln Thr Tyr Gly Asp Thr Asp Ser Val Phe Cys Val His Asn  
 1330 1335 1340  
 Ile Val Gly Asp Gly Gly Met Ile Pro Glu Tyr Asp Glu Gln Thr Gly  
 1345 1350 1355 1360  
 Lys Tyr Tyr Tyr Val Met Asp Ile Ala Leu Lys Asn Lys Met Ala Ala  
 1365 1370 1375  
 Ile Ile Pro Ile Leu Val Asn Ser Leu Thr Lys Gly Ile Gln Phe Val  
 1380 1385 1390  
 Glu Arg Arg Asp Ala Gly Val Gly Met Met Asn Ile Ala His Glu Arg  
 1395 1400 1405  
 Leu Ala Val Ala Gly Leu Leu Phe Ala Lys Lys Thr Tyr His Met Leu  
 1410 1415 1420  
 His Phe Asn Glu Asn Ser Ala Ala Phe Asn Asp Met Ile Lys Leu Lys  
 1425 1430 1435 1440  
 Ser Thr Asp Asn Asn Asn Lys Phe Ala Ser Phe Ile Lys Arg Pro Ser  
 1445 1450 1455  
 His Ala Asp Gly Tyr Val Val Pro His Asn Pro Ser Leu Ile Leu Arg  
 1460 1465 1470  
 Ala Ala Glu Gly Pro Ala Gly Lys Lys Leu Lys Ser Phe Leu Glu Glu  
 1475 1480 1485  
 Glu Gly Ile His Asp Glu Lys Ser Met Glu Glu Trp Phe Thr Ser Ser  
 1490 1495 1500  
 Pro Thr Trp Met Ala Met Asp Ala Ser Val Ile Asn Asn Leu Tyr Ala  
 1505 1510 1515 1520  
 Ser Gln Ile Val Gly Val Glu Lys Gly Asn Trp Ile Asp Ala Met Thr  
 1525 1530 1535  
 Ser Arg Pro Ile Glu Ala Gly Thr Glu Met Met Glu Ala Val Thr Gln  
 1540 1545 1550  
 Ala Asn Ala Ala Phe Thr Pro Tyr Lys Lys Gly Ala Phe Val Lys Lys  
 1555 1560 1565  
 Gly Ile Thr Pro Thr Thr Lys Leu Lys Gly Leu Gln Ser Leu Ile Ala  
 1570 1575 1580  
 Arg Phe Leu Pro Lys Ile Glu Glu Lys Lys Ser Cys Tyr Leu Asp Val  
 1585 1590 1595 1600  
 Met Lys Asn His Val Glu Asn Phe Ala Ser His Ile Thr Asn Pro Ala  
 1605 1610 1615  
 Met Met Ile Thr Ser Ser Arg Val Asn Lys Phe Asp Thr Ser Lys Glu  
 1620 1625 1630  
 Gln Ser Arg Pro Asn Pro Leu Ala Ile Asn Asn His Leu Asn Pro Ser  
 1635 1640 1645  
 Ser Glu Ile Ser Leu Gly Gln Lys Phe Lys Thr Val Thr Ser Val Ser  
 1650 1655 1660  
 Ser Trp Ser Leu Ser Ala Glu Glu Gly Glu Val Pro Ala Gly Tyr Phe  
 1665 1670 1675 1680  
 Asn Ala Gly Ser Val Arg Trp Asp Ala Thr Asn Met Lys Gly Ser Val  
 1685 1690 1695  
 Pro Ala Phe Ser Val Lys Asn Leu Ser Val Val Pro Asn Ala Ile Thr  
 1700 1705 1710  
 Ser Val Tyr Lys Met Val Glu Ser Asp Lys Thr Ala Ile Lys Ser Met  
 1715 1720 1725  
 Ile Ala Lys Asn Val Glu Val Leu Cys Ser Thr Ser Ala Asn Thr Gly  
 1730 1735 1740  
 Phe Ser Leu Arg Arg Gly Ala Leu Ser Phe Asn Thr Gly Val Ile Val  
 1745 1750 1755 1760  
 Thr Lys Asp Val Ala Met Ala Cys Ile Arg Ser Leu Asn Asn Lys Gln  
 1765 1770 1775  
 Met Leu Leu Phe Val Gly Gly Gly Lys Asp Tyr Gly Glu Asp Asp Asp

WO 01/38351

240

PCT/US00/28888

1780	1785	1790
Asp Asp Asp Glu Glu Ala Glu Glu Glu Asp Glu Glu Asn Gly Glu Asn		
1795	1800	1805
Glu Glu Asn Lys Gly Asp Cys Val Thr Glu Lys Lys Ile Pro Gly Arg		
1810	1815	1820
Ser Thr Asn Lys Asp Val Gly Glu Glu Thr Lys Thr Ser Glu Lys Thr		
1825	1830	1835
Glu Gly Glu Arg Lys Gly Ser Lys Thr Ala Lys Gly Lys Thr Glu Glu		
1845	1850	1855
Ile Ala Ser Ser Leu Ser Lys Cys Gly Lys Lys Asp Ala Arg Asp Val		
1860	1865	1870
Ile Leu Asp Arg Leu Leu Lys Ala Thr His Ser Ser Cys Thr Asn Asn		
1875	1880	1885
Glu Glu Arg Thr Arg Val Leu Gln Gln Tyr Ser Asn Cys Thr Leu Ser		
1890	1895	1900
Ser Tyr Ile Thr Ser Val Met Lys Leu Asp Gln Arg Val Ala Asp Gln		
1905	1910	1915
Met Glu Asn Leu Ile Ser Gln Leu Asp Gln Ile Arg Asn Leu Ser Asn		
1925	1930	1935
Lys Lys Arg Gln Glu Lys Gly Gly Pro Phe Lys Ser Glu Leu Asp Ala		
1940	1945	1950
Met Val Ala Ala Val Lys Val Lys Phe Phe Pro Val Leu Asp Ala Ser		
1955	1960	1965
Arg Lys Leu Thr Gln Asp His Trp Lys Lys Cys Pro Val Ser Ile Pro		
1970	1975	1980
Glu Thr Arg Glu Glu Lys Pro Leu Met Gly Val Pro Phe Glu Val Asn		
1985	1990	1995
Ser Leu Ile Gly Lys His Lys Cys Thr Asp Thr Cys Asp Met Ala Cys		
2005	2010	2015
Cys Gln Ser Leu Tyr Phe Val Leu Leu Tyr Thr Leu Ala Leu Lys Phe		
2020	2025	2030
Glu Asn Glu Arg Leu Ala Arg Gln Ile Gly Leu Asp Asp Ser Val Asp		
2035	2040	2045
Leu Met Ala Glu Met Leu Phe Gly Gly Asp Lys Leu Leu Ala Gln Glu		
2050	2055	2060
Val Leu Lys Arg Val Lys Asp Ala Gln Asp Arg Lys Leu Val Lys Ser		
2065	2070	2075
Leu Leu Pro Leu Asn Tyr Asn His Asp Thr Asn Thr Ile Ile Phe Leu		
2085	2090	2095
Phe Glu Ser Leu Arg Phe Ala Gln Lys Pro Val Ala Gly Met Ser Val		
2100	2105	2110
Ser Glu Ile Lys Asp Ala Val Arg Gly Leu Ala Phe Ser Thr Thr Thr		
2115	2120	2125
Gly Thr Val Trp Asn Tyr Thr Asp Glu Arg Phe Phe Gly Pro Leu Tyr		
2130	2135	2140
Asn Met Asp Glu Leu Cys Asn Glu Arg Val Asn Gly Asn Cys Lys Leu		
2145	2150	2155
Ser Phe Ile Thr Gly Ile Tyr His Thr Ala Ala Val Glu Leu Ala Ala		
2165	2170	2175
Ala Cys Leu Ser Cys Val Leu		
2180		

<210> 160

<211> 4172

<212> DNA

<213> SHRIMP

<400> 160

atgactgaac aaggggatca aggaataaaa gtaaggaaat tacatggccc gagaggagaa 60  
agaggagaaa ctgggtccagc aggagcagtt ggccctgcag gccctcaagg agaaagagga 120

gcaattggac	cggcaggaaa	ggatggagca	gttggccctg	caggccctca	aggagaaaga	180
ggagcaattg	gaccggcagg	aaaggatgga	gcagttggcc	ctcaaggccc	tccaggagaa	240
agaggagaaa	atggacgccc	aggaagagat	ggagcagttg	gccctcaagg	agaaagagga	300
gcaattggac	cggcaggaaa	ggatggagca	gttggccctc	aaggagaaag	aggagcaatt	360
ggaccggcag	gaaaggatgg	agcagttggc	cctgcaggcc	ctcaaggaga	aagaggagaa	420
aatggacgcc	caggaagaga	tggagcagtt	ggccctgcag	gccctccagg	agaaaagga	480
gcaattggac	cggcaggaaag	agatggagca	gttggccctg	caggccctcc	aggagaaaga	540
ggagcaacag	gtataccagg	aagggatggc	gtggacggtt	ctgtgggccc	tcaaggagaa	600
agaggagaaa	ttggacgccc	aggaagagat	ggagcagttg	gccctgcagg	ccctcaagga	660
agaagaggag	caacaggacg	cgcaggaaaag	gatggtgcag	ttggtcctgc	aggccctcaa	720
ggagaaaaag	gagaagctgg	taaggacggt	tctatagggc	ctcaagggaat	acaaggccca	780
agaggagaga	ctggaccacc	gggaaggggac	ggcactgcag	cagaaagagg	agaaagaggc	840
ttcccaggac	caccaggcga	aactggacca	ccaggaaagg	atggtgtgga	tggttctgag	900
ggccctcaag	ggaaaagagg	agaaacagga	cccgttggac	ctaggggtga	accaggtcta	960
gctggcctcc	caggaagaga	tggagcaatt	ggccctgcag	gccctccagg	agaaagagga	1020
gcaactggtc	taccagggaag	gaatggtgtg	gatggttcta	tcggccccc	aggaagaaga	1080
ggagcaacag	gccgcgcagg	aaaggatggg	gcagttggcc	ctgcaggccc	tccaggagaa	1140
agaggagcaa	caggtatacc	aggaagggat	ggtgtggacg	gttctgtggg	ccctccagga	1200
gaaagaggag	aaactggacc	agcaggaagg	gacggttcag	ttggccctgc	tggccctcat	1260
ggagaaagag	gagaaaaatg	acgcccagga	agagatgggg	caactggccc	tataggtcct	1320
gctggtcctc	aaggagaaaa	aggagaaaaat	ggacgcccag	gaagagatgg	agcaactggc	1380
cctataggcc	ctagaggaga	aactggtgca	atgggaaaaga	atggcgtgga	cggttctatg	1440
ggtcctcaag	gaagaagagg	agcaacaggc	cgcgcaggaa	aggatggggc	agttggccct	1500
gctggccctc	caggagaaaag	aggagaaaact	ggaccagcag	gaagggacgg	ttcagttggc	1560
cctgctggcc	ctcaaggaga	aacaggatta	actggcagcc	caggaagaga	tggagcaact	1620
ggccctatag	gtcctgctgg	ccctcaagga	gaaaaaggag	aaaatggacg	cccaggaaga	1680
gatggagcaa	ctggccctat	aggtcctgct	ggccctcaag	gagaaaaagg	agaaaaatga	1740
cgccccaggaa	gagatggagc	aactggccct	ataggtcctg	ctggccctca	aggagaaaca	1800
ggattaactg	gacgcccagg	aagagatgga	gcaactggcc	ctataggtcc	tagaggagaa	1860
actggtgcaa	tgggaaagaa	tgggtgtggac	ggttctacgg	gtcctcaagg	aagaagagga	1920
gcaacaggcc	gcgcaggaaa	ggatggagca	gttggccctg	ctggccctcc	aggagaaaga	1980
ggagaaaaatg	gacgcccagg	aagagatgga	gcaactggcc	ctataggtcc	tgctggccct	2040
caaggagaaa	caggattagc	tgggctgcca	ggaagagatg	gagcaatttg	tcctcaagga	2100
gaaaaggagg	aaaatggacg	cccaggaaaag	gatggggcaa	ctggccctat	gggtcctcca	2160
ggagaaaggg	gagagactgg	tcctataggt	cctgctggcc	ctcaaggagc	aactggtctt	2220
ccaggaaggg	atggtgtgga	tgggttctgtt	ggccctcaag	gaaaaagagg	attaataggg	2280
cgcacaggaa	gggatggggc	aattggccct	ttaggtcctg	caggccctaa	aggagaaaca	2340
ggattagctg	gcctggccag	gatagatgga	aaggacggtt	ccgtgggtcc	tcaaggagca	2400
attggaccta	taggcccacg	aggagaaaga	ggagaaactg	gacgaccagg	aagggacggt	2460
gaggatggtt	ccacaggccc	tatgggcccc	caaggactaa	gaggagctac	gggagctcca	2520
ggaccgcaag	gagaaagagg	attaaaggga	cggccaggaa	aagatggtga	aacaggtcct	2580
ccagggcgac	aagggaaggga	tgggaataatg	ggtcctaggg	gtcttcgagg	agaaaaagga	2640
gcacctggta	atgatggtct	agagggacct	gaagggaagag	atggtgcacc	tgggtcccgc	2700
ggccctattg	gacctcaagg	aataagagga	ttaaaaggta	tccagggacg	accaggaaga	2760
gacggagaaa	tgggaccagc	cggcaaggac	ggaatagaag	gccctagagg	tcaagatgga	2820
acaactggcg	ctaaaggacc	tagaggatta	agaggttttc	aaggagaagc	aggagaaact	2880
ggtgcacaa	gatctagagg	agaaaaaggc	gatagagggc	taacaggccc	tcaaggaaga	2940
gacggtccac	ccggtgaaga	aggtcctcaa	ggtccttagag	gagaaaaggg	agcacctggc	3000
cctagaggtc	ctagaggtat	tcgtggccgt	tcaggacctc	aaggaaagta	cggcgtgcaa	3060
ggacctcgag	gtccccgagg	aacaaaagga	agaacaggaa	tacaaggcct	cactggcata	3120
gaaggtcctc	gaggtcctag	aggtatacaa	ggaaagggaag	gaagaatggg	gaaaaattgga	3180
catcgaggag	aaaagggtga	taaaggagac	cgtggagaac	aaggcatcgc	tggagcagac	3240
ggggaaaaag	gtccaaggag	tttacgagga	gtctgaggcc	ctattgggtg	tcctggttaag	3300
cctggcacgg	aaggggttag	aggtcctaga	ggggtgagag	gtgttccttg	ctatcctggc	3360
gcacaagggg	aattaggtcc	ccaaggacca	acaggtcctc	aagggccagc	aggtcctcaa	3420
gggccgatgg	ggcgtacagg	agatactggt	cccatgggcc	ctcctggagc	agtgggacca	3480
agaggagaga	aaggaggtag	aggaagaaag	ggaaaaaatg	gccctaaagg	agcggacgga	3540
aaagatgccg	taaatatcat	acaaaaatat	tcaatcacc	atgctcgtgc	agagataatg	3600
tgggaaggaa	atgaaatcgg	agaagcatac	attggaagat	cttatggaac	tgatacaatc	3660
cctgtgatga	tagaaaaatg	aatagggatg	acaaatgagg	acaaaaaaaa	cgaatattgt	3720
atacaagtaa	tgacaatgca	ctcaataaca	actagaggaa	gaacatcggg	tgttttttgtg	3780

WO 01/38351

PCT/US00/28888

242

gtaagcaata agacagatta taccctttta gttactttac tgatgccaga aagtgtttcc 3840  
 tgtagaacag atgtcagtag aaatgcgagg tcagagaggg tgaatgctgt tagagaaaga 3900  
 gaaagcaaat cgtacagatt tattaggccg tctgaccaat ctataggtag tcattcacgt 3960  
 tcaaaaattg ccgtggtaat gtatccagac gcaagcatga gttactcagt tgatacatta 4020  
 gacgctgatg tggcgcaag agaaacaacg tctgtgcttt tattagcaga aaccatacac 4080  
 ggggaaaaag atagaggttt ctatgctgat agaggaaactg tagggagggt gatggtacct 4140  
 cccactgaag aagagttatt ggtattgcaa gc 4172

<210> 161

<211> 1389

<212> PRT

<213> SHRIMP

<220>

<221> VARIANT

<222> (1)...(1389)

<223> Xaa = Any Amino Acid

<400> 161

Met	Thr	Glu	Gln	Gly	Asp	Gln	Gly	Ile	Lys	Val	Arg	Lys	Leu	His	Gly
1				5					10					15	
Pro	Arg	Gly	Glu	Arg	Gly	Glu	Thr	Gly	Pro	Ala	Gly	Ala	Val	Gly	Pro
			20					25					30		
Ala	Gly	Pro	Gln	Gly	Glu	Arg	Gly	Ala	Ile	Gly	Pro	Ala	Gly	Lys	Asp
		35					40					45			
Gly	Ala	Val	Gly	Pro	Ala	Gly	Pro	Gln	Gly	Glu	Arg	Gly	Ala	Ile	Gly
	50					55					60				
Pro	Ala	Gly	Lys	Asp	Gly	Ala	Val	Gly	Pro	Gln	Gly	Pro	Pro	Gly	Glu
65				70						75				80	
Arg	Gly	Glu	Asn	Gly	Arg	Pro	Gly	Arg	Asp	Gly	Ala	Val	Gly	Pro	Gln
			85						90					95	
Gly	Glu	Arg	Gly	Ala	Ile	Gly	Pro	Ala	Gly	Lys	Asp	Gly	Ala	Val	Gly
			100					105					110		
Pro	Gln	Gly	Glu	Arg	Gly	Ala	Ile	Gly	Pro	Ala	Gly	Lys	Asp	Gly	Ala
		115					120					125			
Val	Gly	Pro	Ala	Gly	Pro	Gln	Gly	Glu	Arg	Gly	Glu	Asn	Gly	Arg	Pro
	130					135					140				
Gly	Arg	Asp	Gly	Ala	Val	Gly	Pro	Ala	Gly	Pro	Pro	Gly	Glu	Arg	Gly
145				150						155				160	
Ala	Ile	Gly	Pro	Ala	Gly	Arg	Asp	Gly	Ala	Val	Gly	Pro	Ala	Gly	Pro
			165						170					175	
Pro	Gly	Glu	Arg	Gly	Ala	Thr	Gly	Ile	Pro	Gly	Arg	Asp	Gly	Val	Asp
		180						185					190		
Gly	Ser	Val	Gly	Pro	Gln	Gly	Glu	Arg	Gly	Glu	Ile	Gly	Arg	Pro	Gly
	195					200					205				
Arg	Asp	Gly	Ala	Val	Gly	Pro	Ala	Gly	Pro	Gln	Gly	Arg	Arg	Gly	Ala
	210					215					220				
Thr	Gly	Arg	Ala	Gly	Lys	Asp	Gly	Ala	Val	Gly	Pro	Ala	Gly	Pro	Gln
225					230					235				240	
Gly	Glu	Lys	Gly	Glu	Ala	Gly	Lys	Asp	Gly	Ser	Ile	Gly	Pro	Gln	Gly
			245						250					255	
Ile	Gln	Gly	Pro	Arg	Gly	Glu	Thr	Gly	Pro	Pro	Gly	Arg	Asp	Gly	Thr
		260						265					270		
Ala	Ala	Glu	Arg	Gly	Glu	Arg	Gly	Phe	Pro	Gly	Pro	Pro	Gly	Glu	Thr
		275					280					285			
Gly	Pro	Pro	Gly	Lys	Asp	Gly	Val	Asp	Gly	Ser	Glu	Gly	Pro	Gln	Gly
	290					295					300				
Lys	Arg	Gly	Glu	Thr	Gly	Pro	Val	Gly	Pro	Arg	Gly	Glu	Pro	Gly	Leu
305					310					315				320	
Ala	Gly	Leu	Pro	Gly	Arg	Asp	Gly	Ala	Ile	Gly	Pro	Ala	Gly	Pro	Pro
			325						330					335	

Gly	Glu	Arg	Gly	Ala	Thr	Gly	Leu	Pro	Gly	Arg	Asn	Gly	Val	Asp	Gly
			340					345					350		
Ser	Ile	Gly	Pro	Gln	Gly	Arg	Arg	Gly	Ala	Thr	Gly	Arg	Ala	Gly	Lys
		355					360					365			
Asp	Gly	Ala	Val	Gly	Pro	Ala	Gly	Pro	Pro	Gly	Glu	Arg	Gly	Ala	Thr
		370				375					380				
Gly	Ile	Pro	Gly	Arg	Asp	Gly	Val	Asp	Gly	Ser	Val	Gly	Pro	Pro	Gly
385					390					395					400
Glu	Arg	Gly	Glu	Thr	Gly	Pro	Ala	Gly	Arg	Asp	Gly	Ser	Val	Gly	Pro
				405					410					415	
Ala	Gly	Pro	His	Gly	Glu	Arg	Gly	Glu	Asn	Gly	Arg	Pro	Gly	Arg	Asp
			420					425					430		
Gly	Ala	Thr	Gly	Pro	Ile	Gly	Pro	Ala	Gly	Pro	Gln	Gly	Glu	Lys	Gly
		435				440					445				
Glu	Asn	Gly	Arg	Pro	Gly	Arg	Asp	Gly	Ala	Thr	Gly	Pro	Ile	Gly	Pro
		450				455					460				
Arg	Gly	Glu	Thr	Gly	Ala	Met	Gly	Lys	Asn	Gly	Val	Asp	Gly	Ser	Met
465					470					475					480
Gly	Pro	Gln	Gly	Arg	Arg	Gly	Ala	Thr	Gly	Arg	Ala	Gly	Lys	Asp	Gly
				485					490					495	
Ala	Val	Gly	Pro	Ala	Gly	Pro	Pro	Gly	Glu	Arg	Gly	Glu	Thr	Gly	Pro
			500					505					510		
Ala	Gly	Arg	Asp	Gly	Ser	Val	Gly	Pro	Ala	Gly	Pro	Gln	Gly	Glu	Thr
		515					520					525			
Gly	Leu	Thr	Gly	Ser	Pro	Gly	Arg	Asp	Gly	Ala	Thr	Gly	Pro	Ile	Gly
	530					535					540				
Pro	Ala	Gly	Pro	Gln	Gly	Glu	Lys	Gly	Glu	Asn	Gly	Arg	Pro	Gly	Arg
545					550					555					560
Asp	Gly	Ala	Thr	Gly	Pro	Ile	Gly	Pro	Ala	Gly	Pro	Gln	Gly	Glu	Lys
				565					570					575	
Gly	Glu	Asn	Gly	Arg	Pro	Gly	Arg	Asp	Gly	Ala	Thr	Gly	Pro	Ile	Gly
		580						585					590		
Pro	Ala	Gly	Pro	Gln	Gly	Glu	Thr	Gly	Leu	Thr	Gly	Arg	Pro	Gly	Arg
		595					600					605			
Asp	Gly	Ala	Thr	Gly	Pro	Ile	Gly	Pro	Arg	Gly	Glu	Thr	Gly	Ala	Met
		610				615					620				
Gly	Lys	Asn	Gly	Val	Asp	Gly	Ser	Thr	Gly	Pro	Gln	Gly	Arg	Arg	Gly
625					630					635					640
Ala	Thr	Gly	Arg	Ala	Gly	Lys	Asp	Gly	Ala	Val	Gly	Pro	Ala	Gly	Pro
				645					650					655	
Pro	Gly	Glu	Arg	Gly	Glu	Asn	Gly	Arg	Pro	Gly	Arg	Asp	Gly	Ala	Thr
			660					665					670		
Gly	Pro	Ile	Gly	Pro	Ala	Gly	Pro	Gln	Gly	Glu	Thr	Gly	Leu	Ala	Gly
		675				680						685			
Leu	Pro	Gly	Arg	Asp	Gly	Ala	Ile	Gly	Pro	Gln	Gly	Glu	Lys	Gly	Glu
	690					695					700				
Asn	Gly	Arg	Pro	Gly	Lys	Asp	Gly	Ala	Thr	Gly	Pro	Met	Gly	Pro	Pro
705					710					715					720
Gly	Glu	Arg	Gly	Glu	Thr	Gly	Pro	Ile	Gly	Pro	Ala	Gly	Pro	Gln	Gly
				725					730					735	
Ala	Thr	Gly	Leu	Pro	Gly	Arg	Asp	Gly	Val	Asp	Gly	Ser	Val	Gly	Pro
			740					745					750		
Gln	Gly	Lys	Arg	Gly	Leu	Ile	Gly	Arg	Thr	Gly	Arg	Asp	Gly	Ala	Ile
		755					760					765			
Gly	Pro	Val	Gly	Pro	Ala	Gly	Pro	Lys	Gly	Glu	Thr	Gly	Leu	Ala	Gly
		770				775						780			
Leu	Pro	Gly	Ile	Asp	Gly	Lys	Asp	Gly	Ser	Val	Gly	Pro	Gln	Gly	Ala
785					790					795					800
Ile	Gly	Pro	Ile	Gly	Pro	Arg	Gly	Glu	Arg	Gly	Glu	Thr	Gly	Arg	Pro
				805					810					815	
Gly	Arg	Asp	Gly	Glu	Asp	Gly	Ser	Thr	Gly	Pro	Met	Gly	Pro	Gln	Gly

WO 01/38351

244

PCT/US00/28888

820	825	830
Leu Arg Gly Ala Thr Gly Ala Pro Gly Pro Gln Gly Glu Arg Gly Leu		
835	840	845
Lys Gly Arg Pro Gly Lys Asp Gly Glu Thr Gly Pro Pro Gly Arg Gln		
850	855	860
Gly Arg Asp Gly Ile Met Gly Pro Arg Gly Leu Arg Gly Glu Lys Gly		
865	870	875
Ala Pro Gly Asn Asp Gly Leu Glu Gly Pro Glu Gly Arg Asp Gly Ala		
885	890	895
Pro Gly Pro Ala Gly Pro Ile Gly Pro Gln Gly Ile Arg Gly Leu Lys		
900	905	910
Gly Ile Gln Gly Arg Pro Gly Arg Asp Gly Glu Met Gly Pro Ala Gly		
915	920	925
Lys Asp Gly Ile Glu Gly Pro Arg Gly Gln Asp Gly Thr Thr Gly Ala		
930	935	940
Lys Gly Pro Arg Gly Leu Arg Gly Phe Gln Gly Arg Thr Gly Glu Thr		
945	950	955
Gly Ala Gln Gly Ser Arg Gly Glu Lys Gly Asp Arg Gly Leu Thr Gly		
965	970	975
Pro Gln Gly Arg Asp Gly Pro Pro Gly Glu Glu Gly Pro Gln Gly Leu		
980	985	990
Arg Gly Glu Arg Gly Ala Pro Gly Pro Arg Gly Pro Arg Gly Ile Arg		
995	1000	1005
Gly Arg Ser Gly Pro Gln Gly Ser Asn Gly Val Gln Gly Pro Arg Gly		
1010	1015	1020
Pro Arg Gly Thr Lys Gly Arg Thr Gly Ile Gln Gly Leu Thr Gly Ile		
1025	1030	1035
Glu Gly Pro Arg Gly Pro Arg Gly Ile Gln Gly Lys Glu Gly Arg Met		
1045	1050	1055
Gly Lys Ile Gly His Arg Gly Glu Lys Gly Asp Lys Gly Asp Arg Gly		
1060	1065	1070
Glu Gln Gly Ile Ala Gly Ala Asp Gly Glu Lys Gly Pro Arg Gly Leu		
1075	1080	1085
Arg Gly Ile Arg Gly Pro Ile Gly Ala Pro Gly Lys Pro Gly Thr Glu		
1090	1095	1100
Gly Val Arg Gly Pro Arg Gly Val Arg Gly Val Pro Gly Tyr Pro Gly		
1105	1110	1115
Ala Gln Gly Glu Leu Gly Pro Gln Gly Pro Thr Gly Pro Gln Gly Pro		
1125	1130	1135
Ala Gly Pro Gln Gly Pro Met Gly Arg Thr Gly Asp Thr Gly Pro Met		
1140	1145	1150
Gly Pro Pro Gly Ala Val Gly Pro Arg Gly Glu Lys Gly Gly Arg Gly		
1155	1160	1165
Arg Lys Gly Lys Asn Gly Pro Lys Gly Ala Asp Gly Lys Asp Ala Val		
1170	1175	1180
Asn Ile Ile Gln Lys Tyr Ser Ile Thr His Ala Arg Ala Glu Ile Met		
1185	1190	1195
Trp Glu Gly Asn Glu Ile Gly Glu Ala Tyr Ile Gly Arg Ser Tyr Gly		
1205	1210	1215
Thr Asp Thr Ile Pro Val Met Ile Glu Asn Arg Ile Gly Met Thr Asn		
1220	1225	1230
Glu Asp Lys Lys Asn Glu Tyr Cys Ile Gln Val Met Thr Met His Ser		
1235	1240	1245
Ile Thr Thr Arg Gly Arg Thr Ser Gly Val Phe Val Val Ser Asn Lys		
1250	1255	1260
Thr Asp Tyr Ile Leu Leu Val Thr Leu Leu Met Pro Glu Ser Val Ser		
1265	1270	1275
Cys Arg Thr Asp Val Ser Thr Asn Ala Arg Ser Val Asn Ala Val Arg		
1285	1290	1295
Glu Arg Glu Ser Lys Ser Tyr Arg Phe Ile Arg Pro Ser Asp Gln Ser		
1300	1305	1310



WO 01/38351

245

PCT/US00/28888

Ile Gly Thr His Ser Arg Ser Lys Ile Ala Val Val Met Tyr Pro Asp  
1315 1320 1325  
Ala Ser Met Ser Tyr Ser Val Asp Thr Leu Asp Ala Asp Val Ala Arg  
1330 1335 1340  
Arg Glu Thr Thr Ser Val Leu Leu Leu Ala Glu Thr Ile His Gly Glu  
1345 1350 1355 1360  
Lys Asp Arg Gly Phe Tyr Ala Asp Arg Gly Thr Val Gly Arg Leu Met  
1365 1370 1375  
Val Pro Pro Thr Glu Glu Glu Leu Leu Val Leu Gln Xaa  
1380 1385

<210> 162  
<211> 201  
<212> DNA  
<213> SHRIMP

<400> 162  
atggccgtaa cagagatacc ttgtggtact cggaacattg cagaagaaga cgtagaatta 60  
gagttgatac ttgtaacagc agaagcagaa gtccgagagg cgatggcagc agctctggcg 120  
gcagcaatca taggggctgt agtggtgcaa atcggcagag tgttgatga agtcgtggca 180  
gcagaggtag agttgatgtg a 201

<210> 163  
<211> 64  
<212> PRT  
<213> SHRIMP

<400> 163  
Met Ala Val Thr Glu Ile Pro Cys Gly Thr Arg Asn Ile Ala Glu Glu  
1 5 10 15  
Asp Val Glu Leu Glu Leu Ile Leu Val Thr Ala Glu Ala Glu Val Arg  
20 25 30  
Glu Ala Met Ala Ala Ala Ala Ala Ile Ile Gly Ala Val Val Val Gln  
35 40 45  
Ile Gly Arg Val Leu Asp Glu Val Val Ala Ala Glu Val Glu Leu Met  
50 55 60

<210> 164  
<211> 627  
<212> DNA  
<213> SHRIMP

<400> 164  
atgcacatgt ggggggttta cgccgctata ctggcggggt tgacattgat actcgtgggt 60  
atatctatag ttgtaaccaa catagaactt aacaagaaat tggacaagaa ggataaagac 120  
gcctaccctg ttgaatctga aataataaac ttgaccatta acggtgttgc tagaggaaac 180  
cactttaact ttgtaaacgg cacattacaa accaggaact atggaaaggt atatgtagct 240  
ggccaaggaa cgtccgattc tgaactggtg aaaaagaaag gagacataat cctcacatct 300  
ttacttggag acggagacca cacactaaat gtaaacaaaag ccgaatctaa agaattagaa 360  
ttgtatgcaa gagtatacaa taatacaaag agggatataa cagtggactc tgtttcactg 420  
tctccaggtc taaatgctac aggaaggga ttttcagcta acaaatttgt attatatttc 480  
aaaccaacag ttttgaagaa aaataggatc aacacacttg tgtttggagc aacgtttgac 540  
gaagacatcg atgatacaaa taggcattat ctgttaagta tgcgattttc tcctggcaat 600  
gatctgttta aggttgggga aaaataa 627

<210> 165  
<211> 204  
<212> PRT  
<213> SHRIMP

<400> 165

Met	His	Met	Trp	Gly	Val	Tyr	Ala	Ala	Ile	Leu	Ala	Gly	Leu	Thr	Leu
1				5					10					15	
Ile	Leu	Val	Val	Ile	Ser	Ile	Val	Val	Thr	Asn	Ile	Glu	Leu	Asn	Lys
			20					25					30		
Lys	Leu	Asp	Lys	Lys	Asp	Lys	Asp	Ala	Tyr	Pro	Val	Glu	Ser	Glu	Ile
		35					40					45			
Ile	Asn	Leu	Thr	Ile	Asn	Gly	Val	Arg	Asn	His	Phe	Asn	Phe	Val	Asn
	50					55					60				
Gly	Thr	Leu	Gln	Thr	Arg	Asn	Tyr	Gly	Lys	Val	Tyr	Val	Ala	Gly	Gln
65					70					75					80
Gly	Thr	Ser	Asp	Ser	Glu	Leu	Val	Lys	Lys	Lys	Gly	Asp	Ile	Ile	Leu
			85					90						95	
Thr	Ser	Leu	Leu	Gly	Asp	Gly	Asp	His	Thr	Leu	Asn	Val	Asn	Lys	Ala
			100					105					110		
Glu	Ser	Lys	Glu	Leu	Glu	Leu	Tyr	Ala	Arg	Val	Tyr	Asn	Asn	Thr	Lys
		115					120					125			
Arg	Asp	Ile	Thr	Val	Asp	Ser	Val	Ser	Leu	Ser	Pro	Gln	Ala	Thr	Gly
	130					135					140				
Arg	Glu	Phe	Ser	Ala	Asn	Lys	Phe	Val	Leu	Tyr	Phe	Lys	Pro	Thr	Val
145					150					155					160
Leu	Lys	Lys	Asn	Arg	Ile	Asn	Thr	Leu	Val	Phe	Gly	Ala	Thr	Phe	Asp
				165				170						175	
Glu	Asp	Ile	Asp	Asp	Thr	Asn	Arg	His	Tyr	Leu	Leu	Ser	Met	Arg	Phe
			180					185					190		
Ser	Pro	Gly	Asn	Asp	Leu	Phe	Lys	Val	Gly	Glu	Lys				
		195					200								

<210> 166

<211> 12543

<212> DNA

<213> SHRIMP

<400> 166

atgtttaagg	caaatgtgct	caacctagga	ggaggcaagt	ttcttgaaag	tgatgtgcga	60
gatacttaaa	taaaatgtgc	caaccagatg	aaagaggaac	caacgactct	acgcatctgt	120
ctctccaaca	aactgcccgga	gtacgacaac	agacgtctac	ctctttctct	tctcaacgaa	180
ggggaacaaa	ttctagtcac	tgataaacttg	accaagaatg	gcaaccctct	ggtcaaacaa	240
atggggcatc	tagctgtaca	agatagagtc	ggcggagacg	gcagtgtcaa	ccccaacaat	300
cttctgtatg	ctggatgtaa	tggttgtaga	tatgacacag	tgaacagagg	aaacgatgga	360
aaattgataa	tgtacagtca	acctgcaact	ttaaaggatg	ttgcaaaatc	caagaagaaa	420
ggcatgtaca	aagtaaaaaa	ggttccagaa	attactgggtg	atcaatttct	ggacaagctg	480
aatgagcggt	cttgccagaa	tgaaaatagg	cgcattggatg	aagaaggacc	ccacgttgga	540
actggaagc	ttctcaggga	acttattatt	atgatgaggc	tttatgagga	agaaacttct	600
agtgacagag	aactatgtgt	gactcctgcc	ttcagagagt	ttctgggttg	tggaagaaca	660
gcaacagacg	tgccagtgtt	taaggttgcc	ttcatcacca	atgcctcttt	gatggggctc	720
aaggtaatat	tctacccaac	catacttgaa	gaggaacgtc	ttgcagctgt	ttcagacaca	780
gaaaatgtgg	tggttgtaaa	gtcaatctta	aaggtaacaac	ttgaattact	gagtgaatgt	840
atgccagaaa	tcgtcgaacg	tggtgaaagt	atgatcaaga	aaactgttgc	ctgtttcaag	900
attgatattg	gcgggagtga	taactggaat	cttcccggac	attgtaaaagt	gagcgacact	960
gcgttcccct	atcaccatgc	gcagttggtg	ggagaaaaga	agaacattct	ctccatcagc	1020
aacgaaaaca	tggtgacttc	tctgggtgtc	gtgaaggcag	atagagccga	agagtggatg	1080
tgcaaaacac	tggaaggttt	cgagaaaaag	tgtctatatc	tcgaaaacct	catgggttct	1140
atggccaata	ccgacgattg	gagaagaaaag	atcctttttt	ctgaacttgg	gccagaaatg	1200
ccttacagaa	acaaaagtgt	gattatggat	caagatttct	gtacaattgg	catgtgttac	1260
aagtttcttg	cagaaggagg	tggttggtta	ttgaccaaga	caaagtgtac	cctattaaag	1320
gagaagatgg	cgtgtaaaag	attggacgat	tctggagatg	gagacgacga	ggaagaagat	1380
aatgaggaag	gaggatctgg	agggaaatct	ggcggcggtt	ctggagacga	gaacaacatc	1440
aacaagcctc	ctccagcccc	taaacaaatc	cctccccttg	ctgccaatgt	ctataatagc	1500

atcatcaatg	acgatgataa	gcttgaccag	attgtgtgtt	tcttcaagag	aaaacacggt	1560
ttctttctga	gtgacattga	caactctcct	cttctggcaa	tggagtttct	tcttccccag	1620
aaagccatga	gcaagaagaa	ttgtgttgaa	cgctgaaac	cagaaacaaa	gaacattatc	1680
aggaatttga	ctggcggtta	caccatcaag	tttgacacaa	tcatgccttt	cgccatcctt	1740
caaatgtgtg	tgagatatga	gaacagaaac	ctgaaattgc	ccagagatac	ggacattcta	1800
caacaaagac	tcaagaataa	cacatgggat	gcactatcaa	agggcaaatt	tgcagaaatg	1860
tggcaattta	cccacaaaga	atcactaaaa	cctcctacaa	ttgaagaatt	ggagagtatt	1920
cctcctccac	ctactcaaag	tgaagaagaa	gctgctgctg	ctgctgctgc	atcaacatca	1980
tcaacaacac	cagatatggt	ttctagtcta	gaagaagggtg	caacatccac	ttcgtcatct	2040
gatgaaaacc	agatagcttc	tttgaaaaat	atcaagaagc	tcttgtccat	catcacgtca	2100
acattcgcca	cgggtgctga	caagaacgac	accatattcg	cgtggacagt	cgttactctt	2160
gctgagcggg	tttgtgcgct	gtacaacatc	acttctcatc	cagaagagta	ctatcagcaa	2220
ataattcgtg	aagattttga	atttgaagga	ggttttgaga	agtttagaca	catgtgcgac	2280
gcaattaacc	gagaacttag	catatatgta	cccaaaagtg	tgcttgaaaa	gcaatctgtc	2340
tgcagaatgg	gagttgcagc	atacgaaaaa	tccatggaga	ggatcaagaa	caagacaaac	2400
agcaaatgtg	gcaaaattaa	gtatgatgaa	agcacaatgg	tctacgaatt	gaacaacgat	2460
acatttaaaa	cgtttgatta	tgatgaatct	gacaagtctt	ttggaccaat	gtatgaatgc	2520
gcccccatgg	agactttcca	acgacttttt	gccagcgtca	agtctgataa	ggaagctgtt	2580
ttggcagaca	agaagtccga	gaagagggag	aaactctatc	aacaaaaaca	agaatatcta	2640
agaaaatgtg	acaacgcagc	cgtttctgct	cgccagatcc	tgaacaatgt	ggccagtaac	2700
ggagcgatg	aagaaagtga	tgaagagagt	gacgatgagg	agaattatgg	agctgccaac	2760
gaaggtgcaa	cagggtgatta	ttatggtggt	gatgatgaag	atgattgtta	tggtttctta	2820
ggcgaatttg	gttctagtga	tgatgaaaat	gtaccatctg	acaatgctag	tagtattaat	2880
aacgtacaag	atgatgtgtt	tagggatgtt	aatttcatca	aaacattcaa	ctttagatcc	2940
agtttatgtc	atcgccaaaa	gtatgtttca	actgtcattg	tggaggaaat	ggagaaaaat	3000
ttgtgtgatg	tactaacatt	ggacaattct	gcagctgaat	ctggtgatat	tttgaaagaa	3060
attaatagaa	gatctttgag	aatgaggaat	tgggtagtgc	cttttactat	gcctgtacga	3120
gaaattgtca	agcctaattg	caatagtga	gacggaaccg	ctaacagcaa	caacaatatc	3180
cctccctttt	gttcttgtgc	ttctcttaat	aacttcaaga	gcgattctcc	ccttagtagt	3240
aacaatacca	tgtctaacga	aaagtgcata	aaactactcc	caattcctag	ttctaacaat	3300
ctgaaagatc	tgactgtagc	tctccgtttc	aacacaatgg	catgtgaaag	gagatacttt	3360
agtgatgtta	ctgcagcttt	aggatttgtg	aaaaaggaca	aagttaatgg	taatattagg	3420
tctatttttag	ataataagag	gtgggatgcc	ataaaacagt	gtaaattggc	aggaaaatgt	3480
ctatcttcag	cattaccttt	gggtatctat	gaaaatgtga	tcagtgaaga	taacaaactc	3540
attaacacgt	ttagacctag	aagtttagcc	agattggcat	gttcatcggg	aggtgatggt	3600
gtatcagata	agagtgttaa	taatggtttc	ttttctggta	tttgggcatt	gtgtgcgaat	3660
caagattgta	agtcgttgtg	gctgggttct	acagtgtgtg	atcctctaaa	accaactaaa	3720
gtgtttaatc	aatcgctttc	tgagaaggaa	cttaaaagaa	aacgtcaaca	gatgtgttta	3780
gatgcagcaa	attattttcaa	ggaccataac	gtttcaaaaac	ttaacatata	cgagtgcctt	3840
aaaatgatgg	aagaatgtat	catgcgtaca	gctttgaatg	gaaaaacaag	taatgattct	3900
gagttctttt	caaatttgat	aacaagatac	ggttctggaa	ccaatagtcc	agcttctaga	3960
ctgtggacaa	ttctagagac	tgtagggag	tgctttaata	acagccttcc	tattgattgg	4020
ggaagtttag	ttaaggattg	ggacggtagt	gacatgctaa	atttaaaggc	aggtgtgtca	4080
aatgtggacg	aatctggcgc	cgtattttgag	ctatcagaat	ttctgggcgt	tagcgcgagg	4140
gcattttttg	gtaaaagatct	ggacacaaat	ttggatgctg	atacgtggga	atgtcttttg	4200
aatgacgaca	ataaggactg	gaaagcgcag	gtcgccaagg	cctatgagtt	tgctctcaaa	4260
gacaacgaca	ttagatctgt	tgaaaatttc	ataaattcaa	gcaacttgct	aacaaataac	4320
aacgtaatta	agaaattgaa	gataaaaacc	acccttagca	atgatgttag	gcatcaaatt	4380
tgggttgaag	acgaatatta	tccacgaaac	aagtctactt	tgcgcagtcg	ggcagagtgg	4440
atggctgcaa	cagaagaggt	tctcaagaca	gaaatgtccc	tctcttgtgt	gctagcaatg	4500
gtcgtatagt	atcgtatcat	gatgcaggga	gaaagtgtgc	gagaaattgc	aactgctcct	4560
ttgcgtctat	ctgtggataa	aatggtacca	ctcattcgat	gctttaaaat	cactagcaag	4620
tggtagcagct	gtacaggcaa	gggagattct	cccaaaaagg	ccgacgctag	cataaaaagaa	4680
ggtagatttt	acgatattga	agaagatcct	ctccactttt	accggttttg	cgcatatggt	4740
attggacaag	ttgcaagtaa	tgatattgtc	attgaagaaa	tgacccgaaa	gattctcatg	4800
tcgtttgatt	ttaatggatt	tgatactagt	aattggttgc	aatttatcac	ctatagattt	4860
tctcacgtcc	taatgggacg	gcgaagcagg	cttctttcta	ggcccccttc	actagttaag	4920
aatttgggta	gtgtttcctc	tttggccgac	aaaaaatagc	aaaaaagcaa	tgatatgtat	4980
gaaaaacgtg	ttgggaaggt	aatgaaacgt	atttcaaggc	ttgttttagt	aaaagcggca	5040
gactctgtgc	gtgcttcttc	aaatgatttg	ttggattgtt	gtattttgga	cgttaatgat	5100
gtcagtgtta	aatctttgga	cgagttccgc	gccaaagact	gcaagagact		5160

agaatcgata	ctaactacaa	tctggttagt	aattcctgta	caactgccca	actcgctgct	5220
gtagaaaagt	cctctagaat	catcaacact	aacatttctt	tccacaacat	tcctgccggt	5280
caagctaaag	taatggacgc	taacgaagaa	gcatttatcg	atccttctct	agaagagatt	5340
aacaaggagg	ataatagtgg	tgcaaaacaa	atgacaggaa	aggggtggaag	taatagggga	5400
cgatcaaaga	aatctggcgg	aggaggtttt	aataacgcgg	gtgggttttta	caatgatgat	5460
tctagtcggg	gatcatcatc	cgtagtgtat	gaggactcga	gatcccgcac	gggtttcagc	5520
cagatacaca	tggacgccag	aaacgaagaa	gacagagaaa	gtggcttggt	ttcgtacgat	5580
ggctatgtac	ttaaccgaat	taagaacatg	atcacacaaa	atcagataaa	taatgacatt	5640
gtgaaagtaa	tttctgacat	tgagaacttt	tttaagattt	gtgttccttt	ctcgaaaaaa	5700
gagtatgctt	tgtacggtgt	cactgaaact	gcgttaagtg	ctggaatgga	tgccattgaa	5760
aggtggaata	aggcagttga	ggaagaaact	aacaagataa	gaaaagagtg	ccgcgatttg	5820
actgatactg	gaagcgttta	tgatatgaac	atcatatgtc	ctggcgatta	catgtcttct	5880
gttggcgagg	gaggaaatgg	tgggtgtggg	ggtggatcat	cctcatcagg	tcaccttctc	5940
tctaataata	ataatgaagc	taaccagacc	aatgaaattt	ctgaagacca	attgaaacat	6000
gaaggaagtg	attgctcggt	ttggtttaac	ttctacaaga	aagtgtgtcaa	ttcttcggaa	6060
aagaagcaag	gaaagggcag	gtttttggca	aatacaggcc	atgagggtag	aatagttggc	6120
cgtcccttga	gaacttttat	ccagtacaag	gggaaagggt	ttgctgaaac	aaaggctctg	6180
acacgctatt	tctccaatca	tatttctgac	tcgtactggg	ctcaagttat	gccaatttgt	6240
tacattaaga	atatggccct	gggagatgag	gacaagtcaa	aaaagaagtt	tggaaagaga	6300
ccctggaaga	actttaacaa	taactctaata	agcagtagta	attcatcggt	gaaatatggt	6360
tcaattcaag	atttggagaa	aaaagactct	ttaaagaatg	ttcctatggg	ttatgatgaa	6420
gattttattgt	ctctttatga	tgaattctta	actacttcca	ctgaaaaatt	agaaaaacatt	6480
aagattgtta	acgattctaa	agacgcgtac	gttatcttgg	gctcttctaa	ccaatcttca	6540
tttgaccaga	ctttctccca	gcaatacttt	acccaccaaa	aaatatccaa	tatcaacacc	6600
tataaaagtt	tgggtaagat	gtggaactgt	aacaatggta	tgtcccctaa	gaaccagata	6660
gttttgctta	aaaaattgct	gtttaaaaat	ctaatacattt	tgtggattaa	attgtacgag	6720
cgtcacatta	gcggttttgg	taactgggga	tgtattcattc	ccaatagttc	aaagaatagc	6780
cactttgaaa	tgacaaagaa	taatgctcca	tgtggcggtta	ccgattcaaa	ccctccactt	6840
tcagtatacc	attctgtgtt	cctgagtggt	gaagactatg	gccaactact	gaaggacact	6900
ttccctctta	tgaatctgca	ccgcactttc	tctgctaaat	ctaaagacaa	taattcctct	6960
gaccttcttc	ccgaaaaaat	ttctgcagct	agtttggtta	aggctgttta	cgcccgtgaa	7020
gttttgcttt	catgttttaga	tcccgaagggt	aatttctgta	cgtcttggtg	cactaatagc	7080
tgttcagtg	tgtttacacc	agggacaaat	ataagacgag	gaggagattt	tttcaacaaa	7140
tcttggtatc	ggcaacaaga	caatgattac	tgtttcattg	gtaaagaaga	aactaagaaa	7200
tgtcccaact	ttgtttcatc	agaaattgaa	attgtcagta	tccttaaaac	tgccgtgttt	7260
ttaagtacca	actctgatgg	ccataagcgc	gtgcttcgtg	tcatacaatta	caataaggac	7320
cattctgggtc	tatatgctgg	cattgatact	ggctgcgcag	acgatgaaga	tgatgatgac	7380
gatcaagggtg	gtacagataa	aacttgccct	ctgcaagaag	attcaatgga	tgctaaggag	7440
atgttaattt	ctatgcgggtc	tgtcattaat	gggaaaagtc	tggacgaatc	tagtctggcc	7500
attaagaagg	ataattttta	cttttttagca	ggaacagata	agggtttcta	cttgataaat	7560
tccttcttta	actcccctgt	gcagggtaaa	tttgtcgctc	ctagaggaac	aaaaatcttt	7620
aagaagtgtt	gcgattttctt	attgaacaaa	ggcacaggag	gagtatttgc	ccgcatattc	7680
tttactgatt	gggcctgtat	tgtttcatca	tcaaagggtg	agaataacaa	gaaggcgatc	7740
gagtctactc	tccaaattag	aaacggggga	tgtttcagcg	agcgattaac	tccttcaatg	7800
tttgacaatg	agtcagaaca	aggcgagttg	ttccacgata	gatactgtcc	cgattttttg	7860
tcagactaca	acaagcagaa	tattttttcg	gaacaagcat	acaaatgctc	atttttagct	7920
aatccagtg	gtcctgcca	gaatatgctc	aaaagagcca	aaaatatacg	tccttgtaatt	7980
acaaatgcag	gaactgctct	catttctaaa	attatggctg	aagttgaaaa	aatgggaaat	8040
gctagaacgt	ttattagtaa	tggtagagca	attcccttca	ggttagctga	aaacactgcc	8100
tgtattagcg	ttgataataa	caggtacttt	ttgattgatg	gtacttatct	tctaggagga	8160
cgcttggaag	ggataaattt	ggtcacggac	atgtatacaa	gggtgaaatt	aaaggcagaa	8220
aaacacgtta	ttttgaaatt	tttattctca	accgaattta	tttctgctgc	tctagcttcg	8280
tccatggaag	gactacaat	gggtagagga	ctttgtttta	tcgaacacgt	atcttatatg	8340
aaaaataccg	actctgtcag	taatatgaac	aaaaactttt	ggtctatggc	tgaagatcag	8400
gaagaaaccg	atgaaaatga	ggatgatgat	gacgaaaatg	aagaagatga	ggatgagaac	8460
gaagaaaaca	cggaaaatac	tagcgtttga	aagtacgaac	ctgtatcaaa	aacagcattc	8520
agttcttctc	taaaacctcc	atcaattttt	attgctgatg	aagattatat	atttttatca	8580
atttttatac	aacttgctaa	agccactagc	gattgtgaaa	cggcttcttc	ctcctcctcc	8640
tcctcctcct	cctcctcttc	ttcttcaaaa	catagtagta	gtagtagtag	tagtaataag	8700
aagaggaagc	aaaaggatga	tgttaattct	acaactacag	cccttcatgc	gctaagaaaa	8760
tgttacatct	catgtgttga	ccaaaaaact	ggcatgccca	gaatggatgt	cgtgtatctt	8820

ttacgtggcc	taatgaactt	tggaggtatg	tgcacagcca	ttgcttcagg	agatggagag	8880
aaagcccacc	acatggttca	aactctttgt	tccgttgcc	tgaatattgc	cacaaaaact	8940
gctgttgtat	ttgtgggaac	aaagggaaac	aattttaaaga	ccactcttgt	cgatttgtgt	9000
aagaggactt	ggttcgaaa	gtttacaaat	attaacgtta	cggctctcaa	taatgctggg	9060
gattcttctt	catctactca	agcaaatgtt	gcttcctttg	ctggcaaaaa	gggtatcgtg	9120
attattgacg	aagttggcca	ccaaggttca	tttggtctaa	agaagtcac	gagcgaagac	9180
gataaggacg	agtctgcttc	cagaagtgg	aatgtcgatt	ttggcggcag	tgggggagaa	9240
atgaatagtg	ttgacatcaa	tgaagcaaga	aacgcgtacg	gtgacggagg	aaactctaag	9300
atcgttttta	gtaatatata	ccgtttaatg	accgaatcaa	aactaaaagt	atgcgaccaa	9360
gagtacgatt	ttatttctga	actcaaacac	gaaaagaata	ggaaaaatgc	ttgtaatgat	9420
acaaaaaaac	gtaaaaagag	tggcgaaatt	gaagatgaag	gggtagaatg	tgaggaaatt	9480
gagagaaaac	acggtaaaaa	tgacgaaaat	ggtgtccgca	taaaggatcc	tatcaatatt	9540
tctttctttg	ccagaaaagc	tcactgggtg	aattgttctt	ctgggtgtgt	ttcaactacc	9600
tttaaagaga	aaaatattgt	gtataatatg	ctccatagag	gcgccatgcc	tttctctatt	9660
aaagactgca	cagattctcc	ttggcttaac	gaaaccgatg	ctgtgtacag	gcactgtaag	9720
aaacctatcg	aatatgaagg	caaattctct	aaatccgaag	taaaaactgc	cttgaagtgc	9780
atattaggtg	aatttggatc	caaaatatgc	gataacgaat	catttgaatc	cattattgat	9840
gaaaactgcc	aggttaataa	tcttcattcc	tggaaacgatt	gtaaagaaga	tatagatgaa	9900
tggaatgaga	aatttatgag	caagaacaag	aaaaataagc	agaatatgaa	aattgaagac	9960
aaagttgatg	cgattatgaa	cataatacag	aaaaacaatg	gccttttgaa	atggaatata	10020
tcatttgatc	gcgacgggtc	gcctgtttta	gtctgcaacc	ctgccactga	aaggttttcc	10080
gagatgatca	cgctcatctt	gtctgcccag	gacatgctcg	aaattaaaaa	atacctgggg	10140
gataattgcc	tgctgactaa	cgggggcggt	aaaaagtcag	ttattgatgg	taatacatcc	10200
gcccctgggt	tattgattgc	gtaccattgc	gtttacactg	gaaaaatttc	tgatgatctt	10260
tcaaaaacaa	actacccagt	tttacttccc	cctcccaaga	aacaacactt	tgtagctgtc	10320
gatgatgctg	ccgagaaggc	tcttctgggc	ctcactctgt	cgaatattaa	tattgacagt	10380
attcgtaaca	tcaagactat	ttctcgcaag	ttatcgacca	taataaagga	ccagaagct	10440
gcaaaactat	tagtggatag	agatctggac	tttatgaaca	tgtacgaacg	ttatgacgca	10500
tctttattcg	atgtgggtta	gaaaccgtca	aaatatagtt	ttcctggatt	tacttctgat	10560
ggttcagtg	ttctgtcaac	ttccacttct	gactgcgaaa	atgtcctttc	ttgcttaaa	10620
aagcgtattg	agaaggataa	gatgtcggtg	aagaattctg	gatcctttat	tcgcatgtgt	10680
atggataaaa	atttgctttc	tgacgaaaag	gacgactctt	cctctaattc	atcaaaaaat	10740
acgtcttctt	tacctaaaac	agacgataac	agcagtgata	ttgcaaaact	tttaagtgtt	10800
tttgagagaa	ataggcaaca	aagttcacaa	ttttcctttg	cttctaactc	tagtgggtga	10860
ggagacagta	ataaagaggc	ttgtttcaat	gtcgatacgc	ctaaaagaag	gcaattgggtg	10920
tctgcactcc	aaaaacataa	cagtgcagga	tcttcttcca	ttattaccga	aatcgctaag	10980
gccatccccc	aaaagaatga	tgtgtccagt	tctataacca	aacatatgct	tccgggcaaa	11040
tttccctcca	gtttattaaa	gaatatgact	tcaccccaaa	acagcgtcat	gattagaggc	11100
attttccaac	agggcgcaaa	aagcagcatt	actgtgtctc	ctattatgat	gtccaacagt	11160
tacatcttct	catctcttgt	cgacgaagct	atgtctaaaa	gacttattgt	ttttccttgt	11220
gatacaacat	ttgtctttga	gaataagaat	gaggatgtta	agaaaattat	cgggcttttg	11280
gatagaggaa	tgaagtatat	coattcatct	ttaatgatgg	aaagatgtat	caaatttggg	11340
aagcatggaa	taaaacagcg	tcaacacgaa	tttaatcatc	acaagaaagc	atggaatgac	11400
tttagtgggc	attccagtga	caacaaaaag	aaggatagaa	ttagtacgt	atcatccgtt	11460
cttccatcag	ttttaatgaa	aaacttgatc	agaacaagg	tcctagaatt	gagagacgta	11520
aaatcagttt	cccgtttaga	ggagaatact	aacacctttt	tccatttcta	cacttcaatg	11580
agtcttttgt	ctaaggctgc	aacaaattct	ggagaatctt	cttcttcttc	tgctactatc	11640
actgaagtag	aagaagacaa	tagtttgtat	gctgaggaac	aacaattgcg	ccgtaagaaa	11700
ccagcaaatt	acgagtcct	gtgcaataaa	ttaccatcac	ctcttcaaat	gtgccagatt	11760
aatcctaatt	ctttgaatac	aatggctatg	aatattgtct	gttctagaca	aggcgcatgg	11820
getcaactaa	atagcatgct	caattcagtt	ttatttgtag	aaatgccctt	tgtaagacc	11880
acaagattct	ttgggagaga	ttttaacatc	aagatgcact	ctcctgccac	aaagaatagg	11940
cagccatca	attttgacaa	ttgtatttgt	atagttttac	caaatcctga	tatggatgtt	12000
gtaggatatg	ataaggaagg	ggaattaata	ggagtgtggt	cttactaac	aaaacatttg	12060
tgtgatgcct	ggggttcaat	ggatgttagg	gatttgatgt	attcttgcca	ccatttacac	12120
atgttgtttg	aaatggctct	ccagtatact	gaatgtaaga	ggagactttc	ctcattaaaa	12180
acactaaaga	gtgataaaa	tgggtgtgat	tatgttgtct	taatgttagc	ctgtatgggtg	12240
taccaattga	tggttagtaa	tctgaaatac	cctgtgtttt	tgctctcttc	atcccacaag	12300
agggcaataa	cagaggatat	tgcggatgag	aatcaagttt	cttactatc	tgttcccatg	12360
tttctggcaa	tggttgtgaa	taaaccgttg	catgctctta	gacactctac	caatttagct	12420
cttctaatg	cttctcaaaa	atctgaccat	tctgatattg	tcaagtacat	tgtaatgaac	12480

caatgggggtt tgagacttaa ccctgattat ctgtgcccaa actgtgtaaa acatgttctt 12540  
tga 12543

<210> 167  
<211> 4152  
<212> PRT  
<213> SHRIMP

<400> 167

Met	Phe	Lys	Ala	Asn	Val	Leu	Asn	Leu	Gly	Gly	Gly	Lys	Phe	Leu	Glu
1				5					10					15	
Ser	Asp	Val	Arg	Asp	His	Leu	Ile	Lys	Cys	Ala	Asn	Gln	Met	Lys	Glu
			20					25				30			
Glu	Pro	Thr	Thr	Leu	Arg	Ile	Cys	Leu	Ser	Asn	Lys	Leu	Pro	Glu	Tyr
		35					40					45			
Asp	Asn	Arg	Arg	Leu	Pro	Leu	Leu	Leu	Asn	Glu	Gly	Glu	Gln	Ile	
	50					55				60					
Leu	Val	Thr	Asp	Asn	Leu	Thr	Lys	Asn	Gly	Asn	Pro	Leu	Val	Lys	Gln
65					70				75						80
Met	Gly	His	Leu	Ala	Val	Gln	Asp	Arg	Val	Gly	Gly	Asp	Gly	Ser	Val
				85					90					95	
Asn	Pro	Asn	Asn	Leu	Leu	Tyr	Ala	Gly	Cys	Asn	Val	Val	Glu	Tyr	Asp
			100					105					110		
Thr	Val	Asn	Arg	Gly	Asn	Asp	Gly	Lys	Leu	Ile	Met	Tyr	Ser	Gln	Pro
		115					120					125			
Ala	Thr	Leu	Lys	Asp	Val	Ala	Lys	Ser	Lys	Lys	Lys	Gly	Met	Tyr	Lys
		130				135					140				
Val	Lys	Lys	Val	Pro	Glu	Ile	Thr	Gly	Asp	Gln	Phe	Leu	Asp	Lys	Leu
145					150					155					160
Asn	Glu	Arg	Ser	Cys	Gln	Asn	Glu	Asn	Arg	Arg	Met	Asp	Glu	Glu	Gly
				165					170					175	
Pro	His	Val	Gly	Thr	Gly	Lys	Leu	Leu	Arg	Glu	Leu	Ile	Ile	Met	Met
			180					185					190		
Arg	Leu	Tyr	Glu	Glu	Glu	Thr	Ser	Ser	Ala	Glu	Lys	Leu	Cys	Val	Thr
		195					200					205			
Pro	Ala	Phe	Arg	Glu	Phe	Leu	Gly	Cys	Gly	Arg	Thr	Ala	Thr	Asp	Val
		210				215					220				
Pro	Val	Phe	Lys	Val	Ala	Phe	Ile	Thr	Asn	Ala	Ser	Leu	Met	Gly	Leu
225					230					235					240
Lys	Val	Ile	Phe	Tyr	Pro	Thr	Ile	Glu	Glu	Arg	Leu	Ala	Ala	Val	Ser
				245					250					255	
Asp	Thr	Glu	Asn	Val	Val	Leu	Leu	Lys	Ser	Ile	Leu	Lys	Val	Gln	Leu
			260					265					270		
Glu	Leu	Leu	Ser	Glu	Cys	Met	Pro	Arg	Ile	Val	Glu	Arg	Val	Glu	Ser
		275					280					285			
Met	Ile	Lys	Lys	Thr	Val	Ala	Cys	Phe	Lys	Ile	Asp	Ile	Gly	Gly	Ser
		290				295					300				
Asp	Asn	Trp	Asn	Leu	Pro	Gly	His	Cys	Lys	Val	Ser	Asp	Thr	Ala	Phe
305					310					315					320
Pro	Tyr	His	His	Ala	Gln	Leu	Val	Gly	Glu	Lys	Lys	Asn	Ile	Leu	Ser
				325					330					335	
Ile	Ser	Asn	Glu	Asn	Met	Val	Thr	Ser	Leu	Gly	Val	Val	Lys	Ala	Asp
			340					345					350		
Arg	Ala	Glu	Glu	Trp	Met	Cys	Lys	Thr	Leu	Glu	Ser	Phe	Glu	Lys	Lys
		355					360					365			
Cys	Leu	Tyr	Leu	Glu	Asn	Leu	Met	Gly	Ser	Met	Ala	Asn	Thr	Asp	Asp
		370				375					380				
Trp	Arg	Arg	Lys	Ile	Leu	Phe	Ser	Glu	Leu	Gly	Pro	Glu	Met	Pro	Tyr
385					390					395					400
Arg	Asn	Lys	Ser	Leu	Ile	Met	Asp	Gln	Asp	Phe	Cys	Thr	Ile	Gly	Met
				405					410					415	

Cys Tyr Lys Phe Leu Ala Glu Gly Gly Gly Leu Leu Leu Thr Lys Thr  
 420 425 430  
 Asn Ala Thr Leu Leu Lys Glu Lys Met Ala Cys Lys Gly Leu Asp Asp  
 435 440 445  
 Ser Gly Asp Gly Asp Asp Glu Glu Asp Asn Glu Glu Gly Gly Ser  
 450 455 460  
 Gly Gly Lys Ser Gly Gly Gly Ser Gly Asp Glu Asn Asn Ile Asn Lys  
 465 470 475 480  
 Pro Pro Pro Ala Pro Lys Gln Ile Pro Pro Leu Ala Ala Asn Val Tyr  
 485 490 495  
 Asn Ser Ile Ile Asn Asp Asp Asp Lys Leu Asp Gln Ile Val Cys Phe  
 500 505 510  
 Phe Lys Arg Lys His Gly Phe Phe Leu Ser Asp Ile Asp Asn Ser Pro  
 515 520 525  
 Leu Leu Ala Met Glu Phe Leu Leu Pro Gln Lys Ala Met Ser Lys Lys  
 530 535 540  
 Asn Cys Val Glu Arg Val Lys Pro Glu Thr Lys Asn Ile Ile Arg Asn  
 545 550 555 560  
 Leu Thr Gly Val Asn Thr Ile Lys Phe Asp Thr Ile Met Pro Phe Ala  
 565 570 575  
 Ile Leu Gln Ile Val Val Arg Tyr Glu Asn Arg Asn Leu Lys Leu Pro  
 580 585 590  
 Arg Asp Thr Asp Ile Leu Gln Gln Arg Leu Lys Asn Asn Thr Trp Asp  
 595 600 605  
 Ala Leu Ser Lys Gly Lys Phe Ala Glu Met Trp Gln Phe Thr His Lys  
 610 615 620  
 Glu Ser Leu Lys Pro Pro Thr Ile Glu Glu Leu Glu Ser Ile Pro Pro  
 625 630 635 640  
 Pro Pro Thr Gln Ser Glu Glu Glu Ala Ala Ala Ala Ala Ala Ser  
 645 650 655  
 Thr Ser Ser Thr Thr Pro Asp Met Val Ser Ser Leu Glu Glu Gly Ala  
 660 665 670  
 Thr Ser Thr Ser Ser Ser Asp Glu Asn Gln Ile Ala Ser Leu Glu Asn  
 675 680 685  
 Ile Lys Lys Leu Leu Ser Ile Ile Thr Ser Thr Phe Ala Thr Gly Ala  
 690 695 700  
 Asp Lys Asn Asp Thr Ile Phe Ala Trp Thr Val Val Thr Leu Ala Glu  
 705 710 715 720  
 Arg Phe Cys Ala Leu Tyr Asn Ile Thr Ser His Pro Glu Glu Tyr Tyr  
 725 730 735  
 Gln Gln Ile Ile Arg Glu Asp Phe Glu Phe Glu Gly Gly Phe Glu Lys  
 740 745 750  
 Phe Arg His Met Cys Asp Ala Ile Asn Arg Glu Leu Ser Ile Tyr Val  
 755 760 765  
 Pro Lys Ser Val Leu Glu Lys Gln Ser Val Cys Arg Met Gly Val Ala  
 770 775 780  
 Ala Tyr Glu Asn Ser Met Glu Arg Ile Lys Asn Lys Thr Asn Ser Lys  
 785 790 795 800  
 Leu Cys Lys Ile Lys Tyr Asp Glu Ser Thr Met Val Tyr Glu Leu Asn  
 805 810 815  
 Asn Asp Thr Phe Lys Thr Phe Asp Tyr Asp Glu Ser Asp Lys Ser Phe  
 820 825 830  
 Gly Pro Met Tyr Glu Cys Ala Pro Met Phe Gln Arg Leu Phe Ala Ser  
 835 840 845  
 Val Lys Ser Asp Lys Glu Ala Val Leu Ala Asp Lys Lys Ser Glu Lys  
 850 855 860  
 Arg Glu Lys Leu Tyr Gln Gln Lys Gln Glu Tyr Leu Arg Lys Cys Asp  
 865 870 875 880  
 Asn Asp Asp Val Ser Ala Arg Gln Ile Leu Asn Asn Val Asn Glu Ser  
 885 890 895  
 Asp Glu Glu Ser Asp Glu Glu Ser Asp Asp Glu Glu Asn Tyr Gly Ala

				900						905				910		
Ala	Lys	Gly	Gly	Ala	Thr	Gly	Asp	Tyr	Tyr	Gly	Gly	Asp	Asp	Glu	Asp	
		915					920					925				
Asp	Cys	Tyr	Gly	Phe	Leu	Gly	Glu	Phe	Gly	Ser	Ser	Asp	Asp	Glu	Asn	
		930				935						940				
Val	Pro	Ser	Asp	Asn	Ala	Ser	Ser	Ile	Asn	Asn	Val	Gln	Asp	Asp	Val	
945				950						955					960	
Phe	Arg	Asp	Val	Asn	Phe	Ile	Lys	Thr	Phe	Asn	Phe	Arg	Ser	Ser	Leu	
				965					970					975		
Cys	His	Arg	Gln	Lys	Tyr	Val	Ser	Thr	Val	Ile	Val	Glu	Glu	Met	Glu	
			980					985					990			
Lys	Asn	Leu	Cys	Asp	Val	Leu	Thr	Leu	Asp	Asn	Ser	Ala	Ala	Glu	Ser	
		995					1000					1005				
Gly	Asp	Ile	Leu	Lys	Glu	Ile	Asn	Arg	Arg	Ser	Leu	Arg	Met	Arg	Asn	
		1010				1015						1020				
Trp	Val	Val	Pro	Phe	Thr	Met	Pro	Val	Arg	Glu	Ile	Val	Lys	Pro	Asn	
1025					1030					1035					1040	
Val	Asn	Ser	Glu	Asp	Gly	Thr	Ala	Asn	Ser	Asn	Asn	Asn	Ile	Pro	Pro	
				1045					1050					1055		
Phe	Cys	Ser	Cys	Ala	Ser	Leu	Asn	Asn	Phe	Lys	Ser	Asp	Ser	Pro	Leu	
			1060					1065					1070			
Ser	Ser	Asn	Asn	Thr	Met	Ser	Asn	Glu	Lys	Cys	Ile	Lys	Leu	Leu	Pro	
		1075					1080					1085				
Ile	Pro	Ser	Ser	Lys	His	Leu	Lys	Asp	Leu	Thr	Val	Arg	Phe	Asn	Thr	
		1090				1095						1100				
Met	Ala	Cys	Glu	Arg	Arg	Tyr	Phe	Ser	Asp	Val	Thr	Ala	Ala	Leu	Gly	
1105				1110						1115					1120	
Phe	Val	Lys	Lys	Asp	Lys	Val	Asn	Gly	Asn	Ile	Arg	Ser	Ile	Leu	Asp	
				1125				1130						1135		
Asn	Lys	Arg	Trp	Asp	Ala	Ile	Lys	Gln	Cys	Lys	Leu	Ala	Gly	Lys	Cys	
			1140					1145					1150			
Leu	Ser	Ser	Ala	Leu	Pro	Leu	Gly	Ile	Tyr	Glu	Asn	Val	Ile	Ser	Glu	
		1155					1160					1165				
Asp	Asn	Lys	Leu	Ile	Asn	Thr	Phe	Arg	Pro	Arg	Ser	Leu	Ala	Arg	Leu	
		1170				1175					1180					
Ala	Cys	Ser	Ser	Gly	Gly	Asp	Gly	Val	Ser	Asp	Lys	Ser	Val	Asn	Asn	
1185				1190						1195					1200	
Gly	Phe	Phe	Ser	Gly	Ile	Trp	Ala	Leu	Cys	Ala	Asn	Gln	Asp	Leu	Glu	
				1205				1210						1215		
Ser	Val	Val	Leu	Gly	Ser	Thr	Val	Val	Asp	Pro	Leu	Lys	Pro	Thr	Lys	
			1220					1225					1230			
Val	Phe	Asn	Gln	Ser	Leu	Ser	Glu	Lys	Glu	Leu	Lys	Glu				



Asp	Asn	Lys	Asp	Trp	Lys	Ala	Gln	Val	Ala	Lys	Ala	Tyr	Glu	Phe	Ala
		1395					1400					1405			
Leu	Lys	Asp	Asn	Asp	Ile	Arg	Ser	Val	Glu	Asn	Phe	Ile	Asn	Ser	Ser
	1410					1415					1420				
Asn	Leu	Leu	Thr	Asn	Asn	Val	Ile	Lys	Lys	Leu	Lys	Ile	Lys	Pro	
1425					1430					1435				1440	
Thr	Pro	Ser	Asn	Asp	Val	Arg	His	Gln	Ile	Trp	Val	Glu	Asp	Glu	Tyr
				1445						1450				1455	
Tyr	Pro	Arg	Asn	Lys	Ser	Thr	Leu	Arg	Ser	Arg	Ala	Glu	Trp	Met	Ala
				1460						1465				1470	
Ala	Thr	Glu	Glu	Val	Leu	Lys	Thr	Glu	Met	Ser	Leu	Ser	Cys	Val	Leu
				1475						1480				1485	
Ala	Met	Val	Ala	Met	Tyr	Arg	Ile	Met	Met	Gln	Gly	Glu	Ser	Val	Arg
	1490					1495					1500				
Glu	Ile	Ala	Thr	Ala	Pro	Leu	Arg	Leu	Ser	Val	Asp	Lys	Met	Val	Pro
1505					1510						1515				1520
Leu	Ile	Arg	Cys	Phe	Lys	Ile	Thr	Ser	Lys	Trp	Cys	Ser	Cys	Thr	Gly
				1525						1530				1535	
Lys	Gly	Asp	Ser	Pro	Lys	Lys	Ala	Asp	Ala	Ser	Ile	Lys	Glu	Gly	Arg
				1540						1545				1550	
Phe	Tyr	Asp	Ile	Glu	Glu	Asp	Pro	Leu	His	Phe	Tyr	Arg	Phe	Ala	Ala
		1555					1560							1565	
Tyr	Val	Ile	Gly	Gln	Val	Asn	Asp	Ile	Val	Ile	Glu	Glu	Met	Thr	Arg
	1570					1575						1580			
Lys	Ile	Leu	Met	Ser	Phe	Asp	Phe	Asn	Gly	Phe	Asp	Thr	Ser	Asn	Trp
1585					1590						1595				1600
Leu	Gln	Phe	Ile	Tyr	Phe	Ser	His	Val	Leu	Met	Gly	Arg	Arg	Ser	Arg
				1605						1610				1615	
Leu	Leu	Ser	Arg	Pro	Leu	Ser	Leu	Val	Lys	Asn	Leu	Val	Ser	Val	Ser
				1620						1625				1630	
Ser	Leu	Ala	Asp	Lys	Asn	Ser	Glu	Lys	Ser	Asn	Asp	Met	Tyr	Glu	Lys
		1635					1640					1645			
Arg	Val	Gly	Lys	Val	Met	Lys	Arg	Ile	Ala	Arg	Leu	Val	Leu	Val	Lys
	1650					1655						1660			
Ala	Ala	Asp	Ser	Val	Arg	Ala	Ser	Ser	Asn	Asp	Leu	Leu	Asp	Cys	Cys
1665					1670						1675				1680
Ile	Leu	Asp	Val	Asn	Asp	Val	Ser	Val	Lys	Ser	Leu	Asp	Glu	Phe	Arg
				1685						1690				1695	
Ala	Lys	Thr	Arg	Gln	Glu	Leu	Gln	Glu	Thr	Arg	Ile	Asp	Thr	Asn	Tyr
				1700						1705				1710	
Asn	Leu	Val	Ser	Asn	Ser	Cys	Thr	Thr	Ala	Gln	Leu	Ala	Ala	Val	Glu
		1715					1720					1725			
Lys	Ser	Ser	Arg	Ile	Ile	Asn	Thr	Asn	Ile	Ser	Phe	His	Asn	Ile	Pro
	1730														

	1875					1880					1885				
Ala	Leu	Tyr	Gly	Val	Thr	Glu	Thr	Ala	Leu	Ser	Ala	Gly	Met	Asp	Ala
	1890					1895					1900				
Ile	Glu	Arg	Trp	Asn	Lys	Ala	Val	Glu	Glu	Glu	Thr	Asn	Lys	Ile	Arg
1905					1910					1915					1920
Lys	Glu	Cys	Arg	Asp	Leu	Thr	Asp	Thr	Gly	Ser	Val	Tyr	Asp	Met	Asn
				1925					1930						1935
Ile	Ile	Cys	Pro	Gly	Asp	Tyr	Met	Ser	Ser	Val	Gly	Glu	Gly	Gly	Asn
			1940					1945						1950	
Gly	Gly	Cys	Gly	Gly	Gly	Ser	Ser	Ser	Ser	Gly	His	Leu	Leu	Ser	Asn
		1955					1960					1965			
Asn	Asn	Asn	Glu	Ala	Asn	Gln	Thr	Asn	Glu	Ile	Ser	Glu	Asp	Gln	Leu
	1970					1975					1980				
Lys	His	Glu	Gly	Ser	Asp	Cys	Ser	Phe	Trp	Phe	Asn	Phe	Tyr	Lys	Lys
1985					1990					1995					2000
Val	Val	Asn	Ser	Ser	Glu	Lys	Lys	Gln	Gly	Lys	Gly	Ser	Val	Leu	Ala
				2005					2010						2015
Asn	Thr	Gly	His	Glu	Gly	Arg	Ile	Val	Gly	Arg	Pro	Leu	Arg	Thr	Phe
			2020					2025						2030	
Ile	Gln	Tyr	Lys	Gly	Lys	Gly	Phe	Ala	Glu	Thr	Lys	Val	Leu	Thr	Arg
		2035					2040					2045			
Tyr	Phe	Ser	Asn	His	Asp	Ser	Tyr	Trp	Ser	Gln	Val	Met	Pro	Ile	Cys
	2050					2055					2060				
Tyr	Ile	Lys	Asn	Met	Ala	Leu	Gly	Asp	Glu	Asp	Lys	Ser	Lys	Lys	Lys
2065					2070					2075					2080
Phe	Gly	Lys	Arg	Pro	Trp	Lys	Asn	Phe	Asn	Asn	Asn	Ser	Asn	Ser	Ser
				2085					2090						2095
Ser	Asn	Ser	Ser	Val	Lys	Tyr	Val	Ser	Ile	Gln	Asp	Leu	Glu	Lys	Lys
			2100					2105				2110			
Asp	Ser	Leu	Lys	Asn	Val	Pro	Met	Gly	Tyr	Asp	Glu	Asp	Leu	Leu	Ser
		2115					2120					2125			
Leu	Tyr	Asp	Asp	Ser	Leu	Thr	Thr	Ser	Thr	Glu	Lys	Leu	Glu	Asn	Ile
	2130					2135					2140				
Lys	Ile	Val	Asn	Asp	Ser	Lys	Asp	Ala	Tyr	Val	Ile	Leu	Gly	Ser	Ser
2145					2150					2155					2160
Asn	Gln	Ser	Ser	Phe	Asp	Gln	Thr	Phe	Ser	Gln	Gln	Tyr	Phe	Thr	His
				2165					2170						2175
Gln	Lys	Ile	Ser	Asn	Ile	Asn	Thr	Tyr	Lys	Ser	Leu	Gly	Lys	Met	Trp
			2180					2185						2190	
Asn	Cys	Asn	Asn	Gly	Met	Ser	Pro	Lys	Asn	Gln	Ile	Val	Leu	Leu	Lys
		2195					2200					2205			
Lys	Leu	Leu	Phe	Lys	Asn	Leu	Asn	Ile	Leu	Trp	Ile	Lys	Leu	Tyr	Glu
	2210					2215					2220				
Arg	His	Val	Leu	Cys	Asn	Trp	Gly	Cys	Ile	His	Pro	Asn	Ser	Ser	Lys
2225					2230					2235					2240
Asn	Ser	His	Phe	Glu	Met	Thr	Lys	Asn	Asn	Ala	Pro	Cys	Gly	Val	

Asn	Asp	Tyr	Cys	Phe	Ile	Gly	Lys	Glu	Glu	Thr	Lys	Lys	Cys	Pro	
2370					2375					2380					
Asn	Phe	Val	Ser	Ser	Glu	Ile	Glu	Ile	Val	Ser	Ile	Leu	Lys	Thr	Ala
2385					2390					2395					2400
Val	Phe	Leu	Ser	Thr	Asn	Ser	Asp	Gly	His	Lys	Arg	Val	Leu	Arg	Val
				2405					2410						2415
Ile	Asn	Tyr	Asn	Lys	Asp	His	Ser	Gly	Ala	Gly	Ile	Asp	Thr	Gly	Cys
			2420					2425					2430		
Ala	Asp	Asp	Glu	Asp	Asp	Asp	Asp	Gln	Gly	Gly	Thr	Asp	Lys	Thr	
		2435					2440					2445			
Cys	Leu	Leu	Gln	Glu	Asp	Ser	Met	Asp	Ala	Lys	Arg	Met	Leu	Ile	Ser
2450					2455						2460				
Met	Arg	Ser	Val	Ile	Asn	Gly	Lys	Ser	Leu	Asp	Glu	Ser	Ser	Leu	Ala
2465					2470					2475					2480
Ile	Lys	Lys	Asp	Asn	Phe	Asn	Phe	Leu	Ala	Gly	Thr	Asp	Lys	Gly	Phe
				2485					2490						2495
Tyr	Leu	Asp	Asn	Ser	Phe	Phe	Asn	Ser	Pro	Val	Gln	Gly	Lys	Phe	Val
			2500					2505					2510		
Ala	Pro	Arg	Gly	Thr	Lys	Ile	Phe	Lys	Lys	Cys	Cys	Asp	Phe	Leu	Leu
		2515					2520					2525			
Asn	Lys	Gly	Thr	Gly	Gly	Val	Phe	Ala	Arg	Ile	Phe	Phe	Thr	Asp	Trp
2530					2535						2540				
Ala	Cys	Ile	Val	Ser	Ser	Ser	Lys	Gly	Lys	Asn	Asn	Lys	Lys	Ala	Ile
2545					2550					2555					2560
Glu	Ser	Thr	Leu	Gln	Ile	Arg	Asn	Gly	Gly	Cys	Phe	Ser	Leu	Thr	Pro
				2565					2570						2575
Ser	Met	Phe	Asp	Asn	Glu	Ser	Glu	Gln	Gly	Glu	Leu	Phe	His	Asp	Arg
			2580					2585					2590		
Tyr	Cys	Pro	Asp	Phe	Leu	Ser	Asp	Tyr	Asn	Lys	Gln	Asn	Ile	Phe	Ser
		2595					2600					2605			
Glu	Gln	Ala	Tyr	Lys	Cys	Ser	Phe	Leu	Ala	Asn	Pro	Val	Cys	Pro	Ala
		2610				2615					2620				
Lys	Asn	Met	Leu	Lys	Arg	Ala	Lys	Asn	Ile	Arg	Leu	Cys	Ile	Thr	Asn
2625					2630					2635					2640
Ala	Gly	Thr	Ala	Leu	Ile	Ser	Lys	Ile	Met	Ala	Glu	Val	Glu	Lys	Met
				2645					2650						2655
Gly	Asn	Ala	Arg	Thr	Phe	Ile	Ser	Asn	Gly	Thr	Ala	Ile	Pro	Phe	Arg
			2660					2665					2670		
Leu	Ala	Glu	Asn	Thr	Ala	Cys	Ile	Ser	Val	Asp	Asn	Asn	Arg	Tyr	Phe
		2675					2680					2685			
Leu	Ile	Asp	Gly	Thr	Tyr	Leu	Leu	Gly	Gly	Arg	Leu	Glu	Gly	Ile	Asn
		2690				2695					2700				
Leu	Val	Thr	Asp	Met	Tyr	Thr	Arg	Cys	Lys	Leu	Lys	Ala	Glu	Lys	His
2705					2710										

2850				2855				2860							
Ser 2865	Lys	His	Ser	Ser 2870	Ser	Ser	Ser	Ser	Ser	Asn 2875	Lys	Lys	Arg	Lys	Gln 2880
Lys	Asp	Asp	Val	Asn 2885	Ser	Thr	Thr	Thr	Ala	Leu 2890	His	Ala	Leu	Arg	Lys
Cys	Tyr	Ile	Ser 2900	Cys	Val	Asp	Gln	Lys 2905	Thr	Gly	Met	Pro	Arg	Met	Asp
Val	Val	Tyr 2915	Leu	Leu	Arg	Gly	Leu	Met 2920	Asn	Phe	Gly	Gly	Met	Cys	Thr
Ala	Ile 2930	Ala	Ser	Gly	Asp	Gly 2935	Glu	Lys	Ala	His 2940	His	Met	Val	Gln	Thr
Leu 2945	Cys	Ser	Val	Asn 2950	Ile	Ala	Thr	Lys	Thr	Ala 2955	Val	Val	Phe	Val	Gly
Thr	Lys	Gly	Asn 2965	Asn	Leu	Lys	Thr	Thr	Leu 2970	Val	Asp	Leu	Cys	Lys	Arg
Thr	Trp	Phe 2980	Glu	Arg	Phe	Thr	Asn 2985	Ile	Asn	Val	Thr	Ala	Leu	Asn	Asn
Ala	Gly 2995	Asp	Ser	Ser	Ser	Ser	Thr 3000	Gln	Ala	Asn	Leu	Ala	Ser	Phe	Ala
Gly 3010	Lys	Lys	Gly	Ile	Val	Ile 3015	Ile	Asp	Glu	Val	Gly 3020	His	Gln	Gly	Ser
Phe 3025	Gly	Ser	Lys	Lys	Ser	Ser	Ser	Glu	Asp 3035	Asp	Lys	Asp	Glu	Ser	Ala
Ser	Arg	Ser	Gly 3045	Asn	Val	Asp	Phe	Gly 3050	Ser	Gly	Gly	Glu	Met	Asn	
Ser	Val	Asp 3060	Ile	Asn	Glu	Ala	Arg	Asn 3065	Ala	Tyr	Gly	Asp	Gly 3070	Gly	Asn
Ser	Lys 3075	Ile	Val	Phe	Ser	Asn	Ile 3080	Asn	Arg	Leu	Met	Thr 3085	Glu	Ser	Lys
Leu 3090	Lys	Val	Cys	Asp	Gln	Glu 3095	Tyr	Asp	Phe	Ile	Ser	Glu 3100	Leu	Lys	His
Glu 3105	Lys	Asn	Arg	Lys	Asn 3110	Ala	Cys	Asn	Asp	Thr 3115	Lys	Lys	Arg	Lys	Arg
Gly	Gly	Glu	Ile 3125	Glu	Asp	Glu	Gly	Val 3130	Glu	Cys	Glu	Glu	Ile	Glu	Arg
Asn	Asp	Gly 3140	Lys	Asn	Asp	Glu	Asn 3145	Gly	Val	Arg	Ile	Lys	Asp 3150	Pro	Ile
Asn	Ile 3155	Ser	Phe	Phe	Ala	Arg	Lys 3160	Ala	His	Trp	Trp	Asn 3165	Cys	Ser	Ser
Gly 3170	Val	Val	Ser	Thr	Thr	Phe 3175	Lys	Glu	Lys	Asn	Ile 3180	Val	Tyr	Asn	Met
Leu 3185	His	Arg	Gly 3190	Ala	Met	Pro	Phe	Ser	Ile	Lys 3195	Asp	Cys	Thr	Asp	Ser
Pro	Trp	Leu	Asn 3205	Glu	Thr	Asp	Ala	Val 3210	Tyr	Arg	His	Cys	Lys	Lys	Pro
Ile	Glu	Tyr 3220	Glu	Gly	Lys	Phe	Ser 3225	Lys	Ser	Glu	Val	Lys	Thr 3230	Ala	Leu
Lys	Cys 3235	Ile	Leu	Gly	Lys	Phe	Gly 3240	Ser	Lys	Ile	Cys	Asp 3245	Asn	Glu	Ser
Phe 3250	Glu	Ser	Ile	Ile	Asp	Glu 3255	Asn	Cys	Gln	Val	Asn 3260	Asn	Leu	His	Ser
Trp 3265	Asn	Asp	Cys 3270	Lys	Glu	Asp	Ile	Asp	Glu	Trp	Asn 3275	Glu	Lys	Phe	Met
Ser	Lys	Asn	Lys 3285	Lys	Asn	Lys	Gln	Asn 3290	Met	Lys	Ile	Glu	Asp	Lys	Val
Asp	Ala	Ile 3300	Met	Asn	Ile	Ile	Gln 3305	Lys	Asn	Asn	Gly	Leu	Leu 3310	Lys	Trp
Asn	Thr 3315	Ser	Phe	Asp	Arg	Asp	Gly 3320	Ser	Pro	Val	Leu	Val 3325	Cys	Asn	Pro
Ala	Thr 3330	Glu	Arg	Phe	Ser	Glu 3335	Met	Ile	Thr	Ser	Ser	Leu	Ser	Ala	Gln

Asp	Met	Leu	Glu	Ile	Lys	Lys	Tyr	Leu	Gly	Asp	Asn	Cys	Leu	Ser	Thr
3345					3350					3355					3360
Asn	Gly	Gly	Val	Lys	Lys	Ser	Val	Ile	Asp	Gly	Asn	Thr	Ser	Ala	Pro
				3365					3370						3375
Gly	Val	Leu	Ile	Ala	Tyr	His	Cys	Val	Tyr	Thr	Gly	Lys	Ile	Ser	Asp
			3380					3385					3390		
Asp	Leu	Ser	Lys	Thr	Asn	Tyr	Pro	Val	Leu	Leu	Pro	Pro	Pro	Lys	Lys
			3395				3400					3405			
Gln	His	Phe	Val	Ala	Val	Asp	Asp	Ala	Ala	Glu	Lys	Ala	Leu	Leu	Gly
	3410					3415				3420					
Pro	Thr	Leu	Ser	Asn	Ile	Asn	Ile	Asp	Ser	Ile	Arg	Asn	Ile	Lys	Thr
3425				3430						3435					3440
Ile	Ser	Arg	Lys	Leu	Ser	Ser	Ile	Ile	Lys	Asp	Pro	Glu	Ala	Ala	Lys
				3445					3450					3455	
Leu	Leu	Val	Asp	Arg	Asp	Leu	Asp	Phe	Met	Asn	Met	Tyr	Glu	Arg	Tyr
			3460					3465					3470		
Asp	Ala	Ser	Leu	Phe	Asp	Val	Val	Lys	Lys	Pro	Ser	Lys	Tyr	Ser	Phe
			3475			3480						3485			
Pro	Gly	Phe	Thr	Ser	Asp	Gly	Ser	Val	Val	Leu	Ser	Thr	Ser	Thr	Ser
	3490				3495					3500					
Asp	Cys	Glu	Asn	Val	Leu	Ser	Cys	Leu	Lys	Lys	Arg	Ile	Glu	Lys	Asp
3505				3510						3515					3520
Lys	Met	Ser	Ala	Lys	Asn	Ser	Gly	Ser	Phe	Ile	Arg	Met	Cys	Met	Asp
				3525					3530					3535	
Lys	Asn	Leu	Leu	Ser	Asp	Glu	Lys	Asp	Asp	Ser	Ser	Ser	Asn	Ser	Ser
			3540					3545					3550		
Lys	Asn	Thr	Ser	Ser	Leu	Pro	Lys	Thr	Asp	Asp	Asn	Ser	Ser	Asp	Ile
			3555			3560						3565			
Ala	Asn	Phe	Leu	Ser	Val	Phe	Gly	Glu	Asn	Arg	Gln	Gln	Ser	Ser	Gln
	3570				3575					3580					
Phe	Ser	Phe	Asn	Ser	Ser	Gly	Gly	Gly	Asp	Ser	Asn	Lys	Glu	Ala	Cys
3585				3590					3595						3600
Phe	Asn	Val	Asp	Thr	Pro	Lys	Arg	Arg	Gln	Leu	Val	Ser	Ala	Leu	Gln
				3605					3610					3615	
Lys	His	Asn	Ser	Asp	Gly	Ser	Ser	Ser	Ile	Ile	Thr	Glu	Ile	Ala	Lys
			3620					3625				3630			
Ala	Ile	Pro	Gln	Lys	Asn	Asp	Val	Ser	Ser	Ser	Ile	Thr	Lys	His	Met
			3635			3640						3645			
Leu	Pro	Gly	Gln	Phe	Pro	Ser	Ser	Leu	Leu	Lys	Asn	Met	Thr	Ser	Pro
	3650				3655					3660					
Gln	Asn	Ser	Val	Met	Ile	Arg	Gly	Ile	Phe	Gln	Gln	Gly	Ala	Lys	Ser
3665				3670					3675						3680
Ser	Ile	Thr	Val	Ser	Pro	Ile	Met	Met	Ser	Asn	Ser	Tyr	Ile	Phe	Ser
				3685											

WO 01/38351

258

PCT/US00/28888

```

3825          3830          3835          3840
Asn Tyr Gly Glu Ser Ser Ser Ser Ser Ala Thr Ile Thr Glu Val Glu
          3845          3850          3855
Glu Asp Asn Ser Cys Asp Ala Glu Glu Gln Gln Leu Arg Arg Lys Lys
          3860          3865          3870
Pro Ala Asn Tyr Glu Ser Met Cys Asn Lys Leu Pro Ser Pro Leu Gln
          3875          3880          3885
Met Cys Gln Ile Asn Pro Lys Ser Leu Asn Thr Met Ala Met Asn Ile
          3890          3895          3900
Ala Arg Ser Arg Gln Gly Ala Trp Ala Gln Leu Asn Ser Met Leu Asn
3905          3910          3915          3920
Ser Val Leu Phe Val Glu Met Pro Phe Val Lys Thr Thr Arg Phe Phe
          3925          3930          3935
Gly Arg Asp Phe Asn Ile Lys Met His Ser Pro Ala Thr Lys Asn Arg
          3940          3945          3950
Pro Ala Ile Asn Phe Asp Asn Cys Ile Gly Met Ser Leu Pro Asn Pro
          3955          3960          3965
Asp Met Asp Val Val Gly Tyr Asp Lys Glu Gly Glu Leu Ile Gly Val
          3970          3975          3980
Gly Ser Ser Leu Thr Lys His Leu Cys Asp Ala Trp Gly Ser Met Asp
3985          3990          3995          4000
Val Arg Asp Leu Met Tyr Ser Cys His His Leu His Met Leu Phe Glu
          4005          4010          4015
Met Ala Leu Gln Tyr Thr Glu Cys Lys Arg Arg Leu Ser Ser Leu Lys
          4020          4025          4030
Thr Leu Lys Ser Asp Lys Thr Gly Val Asp Tyr Val Ala Val Met Leu
          4035          4040          4045
Ala Cys Met Val Tyr Gln Leu Met Val Ser Asn Leu Lys Tyr Pro Val
          4050          4055          4060
Phe Leu Ser Ser Ser Ser His Lys Arg Ala Asn Thr Glu Asp Ile Ala
4065          4070          4075          4080
Asp Glu Asn Gln Val Ser Ser Leu Ser Val Pro Met Phe Leu Ala Met
          4085          4090          4095
Val Val Asn Lys Pro Leu His Ala Leu Arg His Ser Thr Asn Leu Ala
          4100          4105          4110
Leu Pro Asn Ala Ser Gln Lys Ser Asp His Ser Asp Ile Val Lys Tyr
          4115          4120          4125
Ile Val Met Asn Gln Trp Gly Leu Arg Leu Asn Pro Asp Tyr Leu Cys
          4130          4135          4140
Pro Asn Cys Val Lys His Val Leu
4145          4150

```

<210> 168  
 <211> 315  
 <212> DNA  
 <213> SHRIMP

```

<400> 168
atgtacctga gccacattag acaaacacct ttggttgaag aaagacgtgc tctcaccttc 60
aaaatgtacc atcataacaa caataatcag cactcctttg tgaattgccg gtgcaggcga 120
acttcttctc ccatcaactg ttcaagctgt tcccgtgaaa ctttcaactc agtaaaggct 180
atccagtact tcaacaaaac tagcagaaat aatactgcac atcatttcaa gatgccggct 240
tcaaaggatc gcaactattc ttcttttcgag tatgctgaaa cggcagttgc tgctcacaat 300
atatctcagt ggtga                                     315

```

<210> 169  
 <211> 104  
 <212> PRT  
 <213> SHRIMP

atgagcaaga	gttctagttc	tactgtgaaa	tctgcatctt	tttttaactc	tttgatggaa	60
aatgcacctt	cctccaagat	tgaactttta	gaagatggat	ggactaagaa	agctgctgct	120
gctgatacag	atactccaac	tgcaaaacct	actggctctaa	gtatctcgtt	gatggatat	180
agtgggtcaa	tgggaagtg	gaaatcggt	gtggctgata	gttgctctgg	catcatggct	240
acactcaatg	tgattgtctc	aggcattcaa	aatgctatag	tttactacaa	tgatttcgac	300
aaacactcca	ttgaaagtg	acctgttgtg	cgtgcaccag	actgttcaga	gtgggaaggg	360
ggagattttg	tgaagcata	gcgcaagacc	gaagtgtgcg	gaggcggcgg	cggtggcagt	420
gaagctcttc	attcttccct	gatgtacgtg	tttaacaaca	tgattcctgc	attcaagaaa	480
atgcatggca	ttaccaggga	tgaaaaatc	cccactctca	tttttgtgtt	caccgatgag	540
gatgtgcgta	ttgccaatg	tgacatctga	aaattatgtg	ccaactcgta	cgattcggaa	600
actgcaccag	aagaagaatt	gatcatgaag	acttggggac	aaaaaccctt	gactattctt	660
gacatgagga	aggcgctagt	ggagaatgat	tgttggctac	gcatcctcaa	tttctctaga	720
tgtagtggca	gtaaccagag	tgaattgtgc	caagaagatg	tgatcaattt	ttctggatac	780
gacaacaaca	gatggcaact	ttttgaatcc	tttgacagaa	gatcgtgcaa	tgtgcgcaag	840
aatatcgcaa	cctttattat	gaggcaatct	atttccctct	tcaagaacct	gaatgatcaa	900
tttagtgcc	tccccact	gagggaatac	aaccaagaag	aattgaattg	ttttattgaa	960
agtgaaggaa	gatcagagcc	tgcaggattt	gaaaagtatg	gcgacgctca	acgtgaatct	1020
ttcaagtcta	gggttttaaa	tatggcacct	ttagattttg	gaagagttgt	acagggagga	1080
ggaagatata	acaaccacaa	gcgtagcgtg	ttcttgaatt	gtgcatatga	tagtgccctt	1140
tgttgctcta	agcaaacctt	taatccacag	cagcaacaac	aacaacaaca	gtcttctagc	1200
ggtggtggtg	gtatcagcaa	actggctgta	gtgactcaaa	gggccagtc	aattacaggg	1260
ggaggaatg	cgccctctac	tctggcactg	gacatgaacg	cgtgtttcca	gtcgttggac	1320
gatttttgaa	ttgaccatac	aaacctctgt	gattgtaagg	gttgtacaaa	attgatggca	1380
agtgttgaag	ctacatctga	tcaagggaaga	aaaaccaagt	tgtctcgcaa	gtatgctcgc	1440
gtacactggg	ccaagatggt	tgttgaaaaa	ctcttcaaaa	tgatgatcaa	ggaacaatca	1500
atgatgtacg	cttgccagtgc	tgtgcctgat	gaaattgggg	ctattttacg	ctttgttaca	1560
ggcaataatg	caggagtgtg	ctcaggggtc	tctaccatcc	tttctgatct	gggaactgaa	1620
tgcggaacaa	aggcagaata	tgccttccta	aaagaaggaa	agcacatgaa	atctgctagc	1680
tatgatgcac	tccaagtgat	taacaataca	gatctcacac	ctgaacaatc	ctccatgttt	1740
atgtggttct	atgtgcccaa	tgatgccttg	gaagaagctg	gcaaaatctt	ccaccagtca	1800
tttagtttct	ccaattcata	tacaggagga	ggattattat	cactggacga	atacaagagg	1860
ttcgaatttg	gacagtgcct	tgatttcac	agaagaattg	tatcatgtct	aaagattact	1920
agaaatgttg	aagatgtgct	tttggagacg	tcaaaaacct	ccaacagata	ttttgccatc	1980
ccagtctttt	gtggatcaga	tgaccagaag	gaagtgtctc	gagaagaatt	ggccagcgat	2040
ctatttgagg	gacgagaaga	tgtggcagaa	atgatgttta	tcgacttgga	gactgtgatt	2100
caaaaattgg	gaactttgta	tgatgtacct	ctgactctct	cagaaggagg	ttatgcagca	2160
atcaaatctg	tgtgtgcagc	cgccctcttg	gtgctctcgc	tgaagttacc	ctccaacact	2220
tcaaacatga	tctctctctat	tqcaaaaaatg	qcattcacaa	agtattacca	aqagcagaat	2280

<400> 171															
Met	Ser	Lys	Ser	Ser	Ser	Ser	Thr	Val	Lys	Ser	Ala	Ser	Phe	Phe	Asn
1				5					10					15	
Ser	Leu	Met	Glu	Asn	Ala	Pro	Ser	Ser	Lys	Ile	Glu	Leu	Leu	Glu	Asp
			20					25					30		
Gly	Trp	Thr	Lys	Lys	Ala	Ala	Ala	Ala	Asp	Thr	Asp	Thr	Pro	Thr	Ala
		35					40					45			
Lys	Pro	Thr	Gly	Leu	Ser	Ile	Ser	Leu	Met	Asp	Ile	Ser	Gly	Ser	Met
	50					55					60				
Gly	Ser	Val	Lys	Ser	Ala	Val	Ala	Asp	Ser	Cys	Ser	Gly	Ile	Met	Ala
65					70					75					80
Thr	Leu	Asn	Val	Ile	Ala	Pro	Gly	Ile	Gln	Asn	Ala	Ile	Val	Tyr	Tyr
			85						90					95	
Asn	Asp	Phe	Asp	Lys	His	Ser	Ile	Glu	Ser	Gly	Pro	Val	Val	Arg	Ala
			100					105					110		
Pro	Asp	Cys	Ser	Glu	Trp	Glu	Gly	Gly	Asp	Phe	Val	Lys	His	Met	Arg
		115					120					125			
Lys	Thr	Glu	Val	Cys	Gly	Gly	Gly	Gly	Gly	Gly	Ser	Glu	Ala	Leu	His
	130				135						140				
Ser	Ser	Leu	Met	Tyr	Val	Phe	Asn	Asn	Met	Ile	Pro	Ala	Phe	Lys	Lys
145					150					155					160
Met	His	Gly	Ile	Thr	Arg	Asp	Glu	Lys	Phe	Pro	Ile	Leu	Ile	Phe	Val
			165						170					175	
Phe	Thr	Asp	Glu	Asp	Val	Arg	Ile	Ala	Asn	Ser	Asp	Thr	Gly	Lys	Leu
			180					185					190		
Cys	Ala	Asn	Ser	Tyr	Asp	Ser	Glu	Thr	Ala	Pro	Glu	Glu	Glu	Phe	Ile
		195					200					205			
Met	Lys	Thr	Trp	Gly	Gln	Lys	Pro	Leu	Thr	Ile	Leu	Asp	Met	Arg	Lys
	210				215						220				
Ala	Leu	Val	Glu	Asn	Asp	Cys	Trp	Leu	Arg	Ile	Leu	Asn	Phe	Ser	Arg
225					230					235					240



WO 01/38351

PCT/US00/28888

261

Cys	Ser	Gly	Ser	Asn	Gln	Ser	Glu	Leu	Cys	Gln	Glu	Asp	Val	Ile	Asn	
				245					250					255		
Phe	Ser	Gly	Tyr	Asp	Asn	Asn	Arg	Trp	Gln	Leu	Phe	Glu	Ser	Phe	Asp	
			260					265						270		
Arg	Arg	Ser	Cys	Asn	Val	Arg	Lys	Asn	Ile	Ala	Thr	Phe	Ile	Met	Arg	
		275					280					285				
Gln	Ser	Ile	Ser	Leu	Phe	Lys	Asn	Leu	Asn	Asp	Gln	Phe	Ser	Ala	Phe	
	290					295					300					
Pro	Ile	Leu	Arg	Glu	Ile	Asn	Gln	Glu	Glu	Leu	Asn	Val	Phe	Ile	Glu	
305					310					315					320	
Ser	Glu	Gly	Arg	Ser	Glu	Pro	Ala	Gly	Phe	Glu	Lys	Tyr	Gly	Asp	Ala	
				325					330					335		
Gln	Arg	Glu	Ser	Phe	Lys	Ser	Arg	Val	Leu	Asn	Met	Ala	Pro	Leu	Asp	
			340					345					350			
Phe	Gly	Arg	Val	Val	Gln	Gly	Gly	Arg	Tyr	Asn	Asn	His	Lys	Arg		
	355					360					365					
Ser	Val	Phe	Leu	Asn	Cys	Ala	Tyr	Asp	Ser	Ala	Phe	Cys	Cys	Ser	Lys	
	370					375					380					
Gln	Thr	Phe	Asn	Pro	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Ser	Ser	Ser	
385					390					395					400	
Gly	Gly	Gly	Gly	Ile	Ser	Lys	Leu	Ala	Val	Val	Thr	Gln	Arg	Ala	Gln	
				405					410					415		
Ser	Ile	Thr	Gly	Gly	Gly	Asn	Ala	Ala	Ser	Thr	Leu	Ala	Leu	His	Met	
			420					425					430			
Asn	Ala	Cys	Phe	Gln	Ser	Leu	Asp	Asp	Phe	Gly	Ile	Asp	His	Thr	Asn	
	435						440					445				
Leu	Cys	Asp	Cys	Lys	Gly	Cys	Thr	Lys	Leu	Met	Ala	Ser	Val	Glu	Ala	
	450					455					460					
Thr	Ser	Asp	Gln	Gly	Arg	Lys	Thr	Lys	Leu	Ser	Arg	Lys	Tyr	Ala	Arg	
465					470					475					480	
Val	His	Trp	Ala	Lys	Met	Phe	Ala	Glu	Lys	Leu	Phe	Lys	Met	Met	Ile	
				485					490					495		
Lys	Glu	Gln	Ser	Met	Met	Tyr	Ala	Cys	Ser	Ala	Val	Pro	Asp	Glu	Ile	
			500					505					510			
Gly	Ala	Ile	Tyr	Ala	Phe	Val	Thr	Gly	Asn	Asn	Ala	Gly	Val	Cys	Ser	
	515						520					525				
Arg	Val	Ser	Thr	Ile	Leu	Ser	Asp	Leu	Gly	Thr	Glu	Cys	Gly	Asn	Lys	
	530					535					540					
Ala	Glu	Tyr	Ala	Phe	Leu	Lys	Glu	Gly	Lys	His	Met	Lys	Ser	Ala	Ser	
545					550					555					560	
Tyr	Asp	Ala	Leu	Gln	Val	Ile	Asn	Asn	Thr	Asp	Leu	Thr	Pro	Glu	Gln	
				565					570					575		
Ser	Ser	Met	Phe	Met	Trp	Phe	Tyr	Val	Pro	Asn	Asp	Ala	Leu	Glu	Glu	
			580					585					590			
Ala	Gly	Lys	Ile	Phe	His	Gln	Ser	Phe	Ser	Phe	Ser	Asn	Ser	Tyr	Thr	
	595						600					605				
Gly	Gly	Gly	Leu	Leu	Ser	Leu	Asp	Glu	Tyr	Lys	Arg	Phe	Glu	Phe	Gly	
	610					615					620					
Gln	Cys	Phe	Asp	Phe	Ile	Lys	Lys	Leu	Val	Ser	Cys	Leu	Lys	Ile	Thr	
625					630					635					640	
Arg	Asn	Val	Glu	Asp	Val	Leu	Leu	Glu	Thr	Ser	Lys	Thr	Ser	Asn	Arg	
				645					650					655		
Tyr	Phe	Ala	Ile	Pro	Val	Phe	Cys	Gly	Ser	Asp	Asp	Gln	Lys	Glu	Val	
			660					665					670			
Leu	Arg	Glu	Glu	Leu	Ala	Ser	Asp	Leu	Phe	Gly	Gly	Arg	Glu	Asp	Val	
	675						680					685				
Ala	Glu	Met	Met	Phe	Ile	Asp	Leu	Glu	Thr	Val	Ile	Gln	Lys	Leu	Gly	
	690					695					700					
Thr	Leu	Tyr	Asp	Val	Arg	Leu	Ser	Leu	Pro	Glu	Gly	Gly	Tyr	Ala	Ala	
705					710					715					720	
Ile	Lys	Ser	Val	Cys	Ala	Ala	Ala	Ser	Trp	Ala	Ala	Ser	Cys	Glu	Val	

					725					730					735
Pro	Ser	Asn	Thr	Ser	Asn	Met	Ile	Leu	Ser	Ile	Ala	Lys	Met	Ala	Phe
			740					745					750		
Thr	Lys	Tyr	Tyr	Gln	Glu	Gln	Asn	Ser	Ser	Ser	Glu	Thr	Asp	Leu	Asp
		755					760					765			
Ile	Ile	Leu	Pro	Ser	Ile	Gly	Thr	Ala	Asp	Gly	Glu	Ile	Glu	Asn	Asn
	770					775					780				
Leu	Ser	Gly	Val	Val	Phe	Leu	Arg	Cys	Leu	Ile	Thr	Trp	Ala	Asn	Lys
785					790					795					800
Ile	Gly	Val	Asp	Lys	Asn	Phe	Thr	Asn	Lys	Leu	Glu	His	Phe	Leu	Ala
				805					810					815	
Leu	Arg	Ile	Leu	Thr	Lys	Ala	Gly	Asp	Ser	Lys	Ile	Gly	Glu	Lys	Tyr
			820					825					830		
Glu	Thr	Phe	Pro	Val	Arg	Arg	Leu	Asp	Leu	Ser	Glu	Lys	Asp	Leu	Lys
		835					840					845			
Tyr	Ile	Cys	Lys	Arg	Cys	Gly	Val	Lys	Ser	Leu	Lys	Met	Glu	Tyr	Asp
	850					855					860				
Asn	Asp	Glu	Lys	Leu	Cys	Leu	Arg	Cys	Lys	Gly	Asn	Tyr	Arg	Met	Gly
865					870					875					880
Lys	Pro	Met	Val	Tyr	His	Trp	Asp	Asn	Lys	Leu	Thr	Arg	Asp	Pro	Arg
				885					890					895	
Ala	Lys	Thr	Asp	Thr	Thr	Leu	Asn	Leu	Leu	Asn	Ala	Lys	Lys	Ile	Asp
			900					905					910		
Asp	Lys	Val	Lys	Glu	Met	Ala	Ser	Asp	Ile	Ile	Gly	Ala	Leu	Asn	Leu
		915					920					925			
Pro	Pro	Thr	Asp	Lys	Asp	Asn	Glu	Ile	Ala	Val	Ser	Ala	Ala	Ala	Lys
	930					935					940				
Ala	Val	Gly	Ile	Leu	Tyr	Gly	Lys	Thr	Cys	Leu	Leu	Tyr	Lys	Leu	Leu
945					950				955						960
Asn	Glu	Gly	Asn	Ile	Asp	Ile	Pro	Val	Ala	Val	Cys	Val	Glu	Cys	Asp
				965				970						975	
Cys	Cys	Lys	Ser	Lys	Tyr	Met	Met	Ser	Thr	Leu	Gly	Pro	Asp	Lys	Pro
			980					985					990		
Gln	Asn	Arg	Lys	Cys	Pro	Trp	Cys	Arg	Tyr	Ala	Asn	Lys	Leu	Val	Ala
		995					1000					1005			
Met	Gly	Arg	Gly	Gly	Lys	Lys	Leu	Leu	Met	Asp	Leu	Ile	Glu	Cys	Gly
	1010					1015					1020				
Ala	Pro	Ser	Leu	Ala	Met	Val	Glu	Glu	Ala	Ile	Arg	Thr	Ser	Gly	Asp
1025					1030					1035					1040
Val	Met	Tyr	Glu	Glu	Leu	Gly	Glu	Gly	Glu	Glu	Phe	Tyr	Ile	Ile	Asp
				1045				1050						1055	
Tyr	Phe	Leu	Lys	Leu	Lys	Asn	Thr	Ala	Ile	Ala	Glu	Gly	Asn	Lys	Leu
			1060					1065					1070		
Gln	Gln	Asn	Asn	Asn	Lys	Arg	Pro	Ala	Pro	Leu	Gln	Val	Thr	Ser	

WO 01/38351

263

PCT/US00/28888

Val Asn Val Leu Val Arg Gln Lys Ile Cys Val  
1220 1225

<210> 172  
<211> 294  
<212> DNA  
<213> SHRIMP

<400> 172  
atggacattt tggaagacat ctacaagagc gcgatcacgc tcgtactaca atcgcccgaa 60  
tttgtgaatg atgtaaaaca ggaagcttct caggtagttg aggggctaata accttcaatt 120  
agagaagctg tcttttagacg gcttctagaa gaagaaagga aaaaacacga agacgagggtg 180  
ggagatgtgg aagataaaaag acaagcagtg atagacaagg caaatacaat gattacaaca 240  
atggcggcag agtacctgga atctgtagat attttagaag agtttgggtt ttaa 294

<210> 173  
<211> 93  
<212> PRT  
<213> SHRIMP

<400> 173  
Met Asp Ile Asp Ile Tyr Lys Ser Ala Ile Thr Leu Val Leu Gln Ser  
1 5 10 15  
Pro Glu Phe Val Asn Asp Val Lys Gln Glu Ala Ser Gln Val Val Glu  
20 25 30  
Gly Leu Ile Pro Ser Ile Arg Glu Ala Val Phe Arg Arg Leu Leu Glu  
35 40 45  
Glu Glu Arg Lys Lys His Glu Asp Glu Val Gly Asp Val Glu Asp Lys  
50 55 60  
Arg Gln Ala Val Ile Asp Lys Ala Asn Thr Met Ile Thr Thr Met Ala  
65 70 75 80  
Ala Glu Tyr Leu Glu Ser Val Asp Ile Glu Phe Gly Phe  
85 90

<210> 174  
<211> 1530  
<212> DNA  
<213> SHRIMP

<400> 174  
atggcgtccg ggtttgctat caagggaatc gttaaaaact atagaagaat accgtccatt 60  
atcgagtcta tcaaaagtat acggagaagc gaactcgctg aaggtgtata tatcgtctct 120  
ttgcataaaa atactccaaa acatgaggtg gatgaaattg tgaataaaat acgcctctca 180  
gcaggcaacc cctgcttgga aaaaacgtca ttatttcttc aacatcattc acaaatgagg 240  
aatttctata caaggaaagg tgctgaatct gaatctgatt ggctcaaaag actaccagaa 300  
gatttgagga atatcaacaa tatagtgaag agagaagctt taccatga caagtctttc 360  
actttctccc ctctatatag aattctcact gatagactat tcaatgcagc aattcacaac 420  
tgcaagtaca taattgtaac tgctgattta ttgatgggtt gtgggataac caacaacaaa 480  
gtcgaaaaga aactgttaag tatgggtagt attttagggtg gcgaatcaat ggtaccttta 540  
cacgatatat cacatcgatt atcctacaaa ggctccgca tagaaaatcc tatagtgggt 600  
agttgtcatg accaatgctt agttgttcca gtgagtatgt tagggaagat tttttcaagt 660  
aatatgtacc ccacatttaa aaattttgat caatgcatgg cattattttt gaatgcagtt 720  
gttacacatt cggccgaaaa aatggacggg aagcatgaac gtaataagggt catccatatg 780  
ccaaacgagg tataccttga cgccgcaagg aggaaatacc tagaagaaaa actagaggaa 840  
accaacaaat tggatgctat cgatgaagaa gcaagggaag aatatggaaa cgaaatagga 900  
agaataggag acaaaaagtac gtgtctcgtg ttgtcattat ctgcacgaga ctttttcctc 960  
acaaacagat tcaatgaaga cacaccctta tttctgtgta cagaaagagg aatcagattc 1020  
atgtgttcaa attattgtac aatgagagat gaggtgggtt tcaggccccc tttgatcatg 1080  
tctgcctacg ggccaacatc ttaccctatc atcttcaata ctttatatga tcaattcaat 1140

```

gtgcaatatt atccatgtgt ttctggagtt gttttatctt ttattggcga tgatcagtta 1200
gcaccagaac cagaatcatt agtggacatt gttgtacgtt ctataaaaaa tccgtctatt 1260
agaatttttt ctgggtgatgg tgaacacagta taccaggatg gacgtagggt cgaatgttgg 1320
ggtgagggaa agaatcagaa gtttaaccga gaagagcgca ccatttttaa tgtattgagg 1380
ataattaaag catataatga agaacgaact aaagaagatg aagatgaaga ggaagaagag 1440
gaggaagaag aggaacaaca aacagcagca acagtgcacg tagaaagtga ttgggatctc 1500
tcactagaga ggggggagaa ttgggtgtag 1530

```

&lt;210&gt; 175

&lt;211&gt; 507

&lt;212&gt; PRT

&lt;213&gt; SHRIMP

&lt;400&gt; 175

```

Met Ala Ser Gly Phe Ala Ile Lys Gly Ile Val Lys Asn Tyr Arg Arg
1      5      10      15
Ile Pro Ser Ile Ile Glu Ser Ile Lys Ser Ile Arg Arg Ser Glu Leu
20     25     30
Ala Glu Gly Val Tyr Ile Val Ser Leu His Lys Asn Thr Pro Lys His
35     40     45
Glu Val Asp Glu Ile Val Asn Lys Ile Arg Leu Ser Ala Gly Asn Pro
50     55     60
Cys Leu Glu Lys Thr Ser Leu Phe Leu Gln His His Ser Gln Met Arg
65     70     75     80
Asn Phe Tyr Thr Arg Lys Gly Ala Glu Ser Glu Ser Asp Trp Leu Lys
85     90     95
Arg Leu Pro Glu Asp Leu Arg Asn Ile Asn Asn Ile Val Lys Arg Glu
100    105    110
Ala Leu Pro His Asp Lys Ser Phe Thr Phe Ser Pro Leu Tyr Arg Ile
115    120    125
Leu Thr Asp Arg Leu Phe Asn Ala Ala Ile His Asn Cys Lys Tyr Ile
130    135    140
Ile Val Thr Ala Asp Leu Leu Met Gly Cys Gly Ile Thr Asn Asn Lys
145    150    155    160
Val Glu Lys Lys Leu Leu Ser Met Gly Ser Ile Leu Gly Gly Glu Ser
165    170    175
Met Val Pro Leu His Asp Ile Ala His Arg Leu Ser Tyr Lys Gly Leu
180    185    190
Arg Ile Glu Asn Pro Ile Val Gly Ser Cys His Asp Gln Cys Leu Val
195    200    205
Val Pro Val Ser Met Leu Gly Lys Ile Phe Ser Ser Asn Met Tyr Pro
210    215    220
Thr Phe Lys Asn Phe Asp Gln Cys Met Ala Leu Phe Leu Asn Ala Val
225    230    235    240
Val Thr His Ser Ala Glu Lys Met Asp Gly Lys His Glu Arg Asn Lys
245    250    255
Val Ile His Met Pro Asn Glu Val Tyr Leu Asp Ala Ala Arg Arg Lys
260    265    270
Tyr Leu Glu Glu Lys Leu Glu Glu Thr Asn Lys Leu Asp Ala Ile Asp
275    280    285
Glu Glu Ala Arg Glu Glu Tyr Gly Asn Glu Ile Gly Arg Ile Gly Asp
290    295    300
Lys Ser Thr Cys Leu Val Phe Ala Leu Ser Ala Arg Asp Phe Phe Leu
305    310    315    320
Thr Asn Arg Phe Asn Glu Asp Thr Pro Lys Gly Thr Glu Arg Gly Ile
325    330    335
Arg Phe Met Cys Ser Asn Tyr Cys Thr Met Arg Asp Glu Gly Gly Phe
340    345    350
Arg Pro Arg Leu Ile Met Ser Ala Tyr Gly Pro Thr Ser Tyr Pro Ile
355    360    365
Ile Phe Asn Thr Leu Tyr Asp Gln Phe Asn Val Gln Tyr Tyr Pro Cys

```

```

      370      375      380
Val Ser Gly Val Val Leu Ser Phe Ile Gly Asp Asp Gln Leu Ala Pro
385      390      395      400
Glu Pro Glu Ser Leu Val Asp Ile Val Val Arg Ser Ile Lys Asn Pro
      405      410      415
Ser Ile Arg Ile Phe Ser Gly Asp Gly Glu Thr Val Tyr Gln Asp Gly
      420      425      430
Arg Arg Val Asp Val Gly Gly Glu Gly Lys Asn Gln Lys Phe Asn Arg
      435      440      445
Glu Glu Arg Thr Ile Leu Asn Val Leu Arg Ile Ile Lys Ala Tyr Asn
      450      455      460
Glu Glu Arg Thr Lys Glu Asp Glu Asp Glu Glu Glu Glu Glu Glu
465      470      475      480
Glu Glu Glu Gln Gln Thr Ala Ala Thr Val Thr Val Glu Ser Asp Trp
      485      490      495
Asp Leu Ser Leu Glu Arg Gly Glu Asn Trp Val
      500      505

```

<210> 176  
 <211> 246  
 <212> DNA  
 <213> SHRIMP

```

<400> 176
atgacttgctc cagaaatctc taaacacatt tctggaacag acagacgttt ctggaacacg 60
gctgacccag gtggcctcag ctatcctttc aaccctcttt ttacccttca tctccatctc 120
aaaaactttt caaaaatttt ttcagctcac tccagtttag ggggtggacc gctgactagg 180
ccttatgtca agttcgaagg gtggaccgct gggctcgaccc aacgtcagat tacagagagg 240
agctag 246

```

<210> 177  
 <211> 77  
 <212> PRT  
 <213> SHRIMP

```

<400> 177
Met Thr Cys Pro Glu Ile Ser Lys His Gly Thr Asp Arg Arg Phe Trp
1      5      10      15
Asn Thr Ala Asp Pro Gly Gly Leu Ser Tyr Pro Phe Asn Pro Leu Phe
      20      25      30
Thr Leu His Leu His Leu Lys Asn Phe Ser Lys Ile Phe Ser Ala His
      35      40      45
Ser Ser Leu Gly Gly Gly Pro Leu Trp Tyr Val Lys Phe Glu Gly Trp
      50      55      60
Thr Ala Gly Ser Thr Gln Arg Gln Ile Thr Glu Arg Ser
65      70      75

```

<210> 178  
 <211> 738  
 <212> DNA  
 <213> SHRIMP

```

<400> 178
atggtttcca ccaggtctat ggaagcaaaa gctgcagcag cagcaaaagc aaaagaagtt 60
tctcccacga ccagtaagag aaaggcggag gacctcactg aaggaacaga agaagaagaa 120
gaatcagtag aaacacaccc gccgagtaag ctcccagag tcgatgaaga tgaagtctat 180
attgatgaaa atgttgatgg tgatgtgcag atcctcgccct catcaatcga agtcgccaga 240
atggagagag aaagacttgc cgaagccatg gtccgagaca taaaaatcga ggaagaaaaa 300
gccgcaacgg aagcgaggaa agaaatagcc tctcgccctaa ttataaaga aatggtatat 360

```

cttttgccctc	aactggaaaa	catgactaac	cgcctccgtc	cgagatcact	tctcaggcac	420
aacgaaatga	ccattacaga	ccgcacgttc	agtgatttgc	agatattcaa	caaagtcact	480
tttgaattcc	ctatactgac	tgatattgct	ttccttgccc	gtgaaaaatc	acgtgtcgag	540
ggttcgagat	tctacaacga	tatgaagatt	ggacctataa	cagcctacaa	attgaatttg	600
atgtgtaata	aattcataga	gtctgtttgt	caaaagggtga	aggcagaaat	atccccattt	660
gttgaaggta	tggtcatcaag	tgaacttgaa	gggtcacctt	tttgggattt	caagcaaaga	720
atagtaaaac	acacctag					738

```
<210> 179
<211> 245
<212> PRT
<213> SHRIMP
```

[illegible]

<210>	180
<211>	1221
<212>	DNA
<213>	SHRIMP

```
<400> 180
atgtctcaca tcaactctac ctctgctgcc acgacttcat ccaacactct gccgatttgc 60
accactacag cccctatgat tgctgccgcc agagctgctg ccatgcctc tcggacttct 120
gctctgctg ttacaagtat caactcta tctacgtctt cttctgcaat gttccgagta 180
ccacaaggta tctctgttac ggccatgcct cccgtgccag cacttacatc tctgactgaa 240
tctaactgaa cgaggatgtc ttctacaccc aatgtggatg ttatacctgt tcttggcccc 300
aagaacaaag ccaagtctaa gaagaaggat tcaaaaggag agaagaacca gaatggcaac 360
cgtagcaqtg acgaqqacga accatctctt qttatcqacg acqgttcttg aagacaqtct 420
```

```

aagaacaaga aatattcttg ggtcacatct cttgctacta ctacggctga aagaaacaac 480
gacactctcg cccacacctag gcccttcctt cccacacccg aagaaggaaa tatgtctgaa 540
attgacgcag ggctaagtaa tccagtcact cgccaaatca ccggagaagt ttatagcgct 600
gcactcactt ctggagttgg agataatgga ctatatcctt cccacttcac ggttgctgac 660
acttcttacg gagattgcga aacacccata cctggacctg cttttgtcct cgacgacggg 720
acagttagca gaggcacatc tcttctgcac agagaagagg cagaattctt gaatgatgga 780
agtaagggtga tccataccgt taaaccaaga aacagcaagt actccaatat tcaacgtgcc 840
gctagctgta tggcctacgc tgtggacctt ctaaacaacc ataatatcac ctctgaccaa 900
tttgatttta tggctatgac tgcattgggca gcccgtaac gttgtggaga aatggccaag 960
ttttttgaga agcgcgataa ggacatcgga gaatatagga ataagggtgtt ccaataacaac 1020
agaggcatct ttacacgcac cactgaaatg aataaacgcg caaagattat cctggaacaa 1080
caacaacgcc gtgaagctgc tgccgctgcc gctgccaccg gtgccaccgc ccctatccct 1140
acaacttctg ctgccggagt tgggtgctact tcttctgcta ctactaactc tctcgaatat 1200
caagaaatca gataccagta a                                     1221

```

<210> 181  
 <211> 402  
 <212> PRT  
 <213> SHRIMP

<400> 181

Met	Ser	His	Ile	Asn	Ser	Thr	Ser	Ala	Ala	Thr	Thr	Ser	Ser	Asn	Thr
1				5					10					15	
Leu	Pro	Ile	Cys	Thr	Thr	Thr	Ala	Pro	Met	Ile	Ala	Ala	Ala	Arg	Ala
			20					25					30		
Ala	Ala	Ile	Ala	Ser	Arg	Thr	Ser	Ala	Ser	Ala	Val	Thr	Ser	Ile	Asn
		35					40					45			
Ser	Asn	Ser	Thr	Ser	Ser	Ser	Ala	Met	Phe	Arg	Val	Pro	Gln	Gly	Ile
	50					55					60				
Ser	Val	Thr	Ala	Met	Pro	Pro	Val	Pro	Ala	Leu	Thr	Ser	Leu	Thr	Glu
65					70					75				80	
Ser	Thr	Gly	Thr	Arg	Met	Ser	Ser	Thr	Pro	Asn	Val	Asp	Val	Ile	Pro
				85					90					95	
Val	Pro	Gly	Pro	Lys	Asn	Lys	Ser	Lys	Ser	Lys	Lys	Lys	Asp	Ser	Lys
			100					105					110		
Arg	Lys	Lys	Asn	Gln	Asn	Gly	Asn	Arg	Ser	Ser	Asp	Glu	Asp	Glu	Pro
			115				120					125			
Ser	Leu	Val	Ile	Asp	Asp	Gly	Ser	Gly	Arg	Gln	Ser	Lys	Asn	Lys	Lys
						135					140				
Tyr	Ser	Trp	Val	Thr	Ser	Leu	Ala	Thr	Thr	Thr	Ala	Glu	Arg	Asn	Asn
145					150					155				160	
Asp	Thr	Leu	Ala	Pro	Pro	Arg	Pro	Phe	Leu	Pro	Thr	Pro	Glu	Glu	Gly
				165					170					175	
Asn	Met	Ser	Glu	Ile	Asp	Ala	Gly	Leu	Ser	Asn	Pro	Val	Thr	Arg	Gln
			180					185					190		
Ile	Thr	Gly	Glu	Val	Tyr	Ser	Ala	Leu	Thr	Ser	Gly	Val	Gly	Asp	
			195				200					205			
Asn	Gly	Pro	Ser	His	Phe	Thr	Val	Ala	Asp	Thr	Ser	Tyr	Gly	Asp	Cys
						215					220				
Glu	Thr	Pro	Ile	Pro	Gly	Pro	Ala	Phe	Val	Leu	Asp	Asp	Gly	Thr	Val
225					230					235				240	
Ser	Arg	Gly	Thr	Ser	Leu	Leu	His	Arg	Glu	Ala	Glu	Phe	Leu	Asn	
				245					250				255		
Asp	Gly	Ser	Lys	Val	Ile	His	Thr	Val	Lys	Pro	Arg	Asn	Ser	Lys	Tyr
			260					265					270		
Ser	Asn	Ile	Gln	Arg	Ala	Ala	Ser	Cys	Met	Ala	Tyr	Ala	Val	Asp	Leu
			275				280					285			
Leu	Asn	His	Asn	Ile	Thr	Ser	Asp	Gln	Phe	Asp	Phe	Met	Ala	Met	
			290			295				300					
Thr	Ala	Trp	Ala	Ala	Arg	Gln	Arg	Cys	Gly	Glu	Met	Ala	Lys	Phe	Phe
305					310					315					320

WO 01/38351

PCT/US00/28888

268

Glu	Lys	Arg	Asp	Lys	Asp	Ile	Gly	Glu	Tyr	Arg	Asn	Lys	Val	Val	Gln
				325					330					335	
Tyr	Asn	Arg	Gly	Ile	Phe	Thr	Arg	Thr	Thr	Glu	Met	Asn	Lys	Arg	Ala
			340					345					350		
Lys	Ile	Ile	Gln	Gln	Gln	Arg	Arg	Glu	Ala	Ala	Ala	Ala	Ala	Ala	Ala
		355					360					365			
Thr	Gly	Ala	Thr	Ala	Pro	Ile	Pro	Thr	Thr	Ser	Ala	Ala	Gly	Val	Gly
	370					375					380				
Ala	Thr	Ser	Ser	Ala	Thr	Thr	Asn	Ser	Leu	Glu	Tyr	Gln	Glu	Ile	Arg
385					390					395					400
Tyr	Gln														

<210> 182  
 <211> 1617  
 <212> DNA  
 <213> SHRIMP

<400> 182  
 atggaagact ttaaacaatt aaaagtaaaa aatgggtatgt gtttgctctgg ggaaaatacc 60  
 gaaaattatg aacgggtact attaacattc aaatcagtc agagtgtcag gagaagtggag 120  
 ctaaaggaag gacattttat agttcgtctt agagacaagg aagtactcca catcaagaac 180  
 ggtaacgaaa gattgagaca attaacagga gatcctacgc ttcagattgg actaaaatac 240  
 acatccagtc tcccaaaaaca aggtagtttc ttagaagatg aagaccctaa ttatggaaaa 300  
 aaatggaacg aatcactacc aagcccattc caggaaatga acaaaattgt ggaagaaaag 360  
 gctctagtta atgacaagaa ctttaaattt tcacccctat acagaatcat acatgaacgt 420  
 ctttcaaatt cgcccggtta gaaatgtgat tatatgataa tcacaacaga cttcttagta 480  
 ggggtgtgggt tttctcctag aaattgtacc cgtactctta agaatatgga acaagtgtta 540  
 gtgcaacacg gtggtacctc ttctcgtgta tcagtgtatg atatctgtga taggttaacg 600  
 tacaatggct taagtatcgc aaaccccata gttggcagtt tttcaaatat gtgcctaatt 660  
 gtaccaatgg ataaacttgg attacttttc tacaacagca cacaccgctc agctaaaagc 720  
 attggaatatt acatgtcatg ccttttcaat gctgcagttg tatacacgct agaaaagagt 780  
 aatcaaaaat tagataattt cgaaaaggaa atcagatttg caaaaaatga agtcaacctt 840  
 ctagttagcg aaagaagtgt tctggaagaa aaacttaaa aatccaaaaa gctatatgct 900  
 gcctcagaag aacaaaggat ttctcttcga gatgtgcata aaaagtcctc aattgcatca 960  
 tccagatatg acggcggtgc ctgtctggtc tttgcctttt ctgaccgaga tttctccttg 1020  
 ttgtgcagaa ccaatggaaa tggttccttt tactctgcca cagaagaagg aatcagatac 1080  
 gtctcttcgg acgactacag aaagagggac gtggatgaac gtaggcccag attggctcatg 1140  
 tccataactg gctcagatgc acctatatgc atcagagata gtatacgaaa ccattttaat 1200  
 aaccatttca ttgcatccgg aaagggtaat gaaatatcat tcatcgatcc tccgaatgaa 1260  
 aggtttgttg tggatgggt cagagaggtt actggatcag acatcaaaat cttcatggat 1320  
 aatggaaaag tatatcaaga tgggtgtagaa ataaaagtga ttgacccctc ttctaagaa 1380  
 ggcaaggaca taataaaaaa ggaagaaaca ttaccagagg aggaaaggaa gcgtctgcgc 1440  
 cgagagcgct gcatgatttt caacacagtt aaggcaattg agacgtacaa cgaggaaacgt 1500  
 ggggaagaag aagaagtagc cacaagcagt ggaggaacaa agagaaagag ggaggagaaa 1560  
 gaaggcgatt atgttgcctt tttgaacaag gcatgcaaa aatttaaagt ttgttga 1617

<210> 183  
 <211> 534  
 <212> PRT  
 <213> SHRIMP

<400> 183  
 Met Glu Asp Phe Lys Gln Leu Lys Val Lys Asn Gly Ile Cys Leu Ser  
 1 5 10 15  
 Gly Glu Asn Thr Glu Asn Tyr Glu Arg Val Leu Leu Thr Phe Lys Ser  
 20 25 30  
 Val Lys Ser Val Arg Arg Ser Glu Leu Lys Glu Gly His Phe Ile Val  
 35 40 45  
 Arg Leu Arg Asp Lys Glu Val Leu His Ile Lys Asn Gly Asn Glu Arg



[illegible]

WO 01/38351

270

PCT/US00/28888

<210> 184  
<211> 1386  
<212> DNA  
<213> SHRIMP

<400> 184  
atggactcat ctgcatctgt cgtgtttatg agattcgccc ctcccgggga ggaaactgca 60  
cttccgcccc gacgtgccac gcccggttct gtgcctacg acctatttcc ctctgaagaa 120  
atggatatcg aacctatggg actggccaag atctctactg gatattggaat agacaagttt 180  
cccgcaggct gttatggaca aattgtgtca cgttctggga tgacatggaa gaacaacact 240  
agtgtaccta ctggaacgat tgatgtggat tataggggag aattgaaagt gattctgcgc 300  
aaccatagtg cagaaaaaag tgtccaatc agaaagggaa ccagcattgc ccagttgatt 360  
ttcttaagat attgtgatgt cgaggaagaa cagattgtgt atattaatga aaccacggga 420  
gagagaacga ttattgactc tagttctaaa aaggacaaca aaaatcaagc aagaagcgtg 480  
cgtggaactg gtggatttgg atctacagat aacccaaatt ttactgaaac caccgtctca 540  
agaaaccaac aagaagagaa caaaaaggaa gaattggaag aaggggagat cgtagaaatg 600  
gaaggtttta ttgacattcc ttttcttgaa ggtttcgaaa ataccctcgc agaacaaagc 660  
aacgaaactg gtgtgacata ccctaatacg aatcaagatg tggaagaaaa agatactaaa 720  
aatatagatg tcgtcagaga attggaagct gaatttagta gtggaattgg gagtggctcc 780  
atggactctt ctgactcadc cgattcttct tcttcttctc ctgactcadc cgattcgtct 840  
gattcatctg actctgaatc atctgatgat tcagaaggag gggataataa ggtccgaaga 900  
ataagacgtc atcagtatca ccggcgccag ttgagttatt cggatgacgt caatggaggg 960  
ggaagaaatt ctgagaaaaa ggagatggac agagtaactc acataaaaac tgaacacata 1020  
aaaagagagg acgaaccag atacgaagaa agagaaagat atattcatcc aagaagaatg 1080  
caagtgccca aggactatta ttgtgagcaa tacgaacact acgacgcccc tgctgctgct 1140  
caccaccacc gccaccacca acaccgccac caacaccaga ggcactttaa ccaaccccg 1200  
tccaacaatt cttctgacgt tactgcttac gtcaatgaaa attccccac gaggccatgc 1260  
cgtgatcgca actctcgatt ctcagaaaga cccaacaatg gcggttataa ccggatcaac 1320  
tcaaggtata caactttcga cccttataga tatggcgcaa gaagagggcg tggaggagta 1380  
tattag 1386

<210> 185  
<211> 457  
<212> PRT  
<213> SHRIMP

<400> 185  
Met Asp Ser Ser Ala Ser Val Val Phe Met Arg Phe Ala Pro Pro Gly  
1 5 10 15  
Glu Glu Thr Ala Leu Pro Pro Arg Arg Ala Thr Pro Gly Ser Val Ala  
20 25 30  
Tyr Asp Leu Phe Pro Ser Glu Glu Met Asp Ile Glu Pro Met Gly Leu  
35 40 45  
Ala Lys Ile Ser Thr Gly Tyr Gly Ile Asp Lys Phe Pro Asp Gly Cys  
50 55 60  
Tyr Gly Gln Ile Val Ser Arg Ser Gly Met Thr Trp Lys Asn Asn Thr  
65 70 75 80  
Ser Val Pro Thr Gly Thr Ile Asp Val Asp Tyr Arg Gly Glu Leu Lys  
85 90 95  
Val Ile Leu Arg Asn His Ser Ala Glu Lys Ser Val Pro Ile Arg Lys  
100 105 110  
Gly Thr Ser Ile Ala Gln Leu Ile Phe Leu Arg Tyr Cys Asp Val Glu  
115 120 125  
Glu Glu Gln Ile Val Tyr Ile Asn Glu Thr Thr Gly Glu Arg Thr Ile  
130 135 140  
Ile Asp Ser Ser Ser Lys Lys Asp Asn Lys Asn Gln Ala Arg Ser Val  
145 150 155 160  
Arg Gly Thr Gly Gly Phe Gly Ser Thr Asp Asn Pro Asn Phe Thr Glu  
165 170 175

WO 01/38351

PCT/US00/28888

271

Thr Thr Val Ser Arg Asn Gln Gln Glu Glu Asn Lys Lys Glu Glu Leu  
180 185 190  
Glu Glu Gly Glu Ile Val Glu Met Glu Gly Phe Ile Asp Ile Pro Phe  
195 200 205  
Leu Glu Gly Phe Glu Asn Ile Leu Ala Glu Gln Ser Asn Glu Thr Gly  
210 215 220  
Val Thr Tyr Pro Asn Thr Asn Gln Asp Val Glu Glu Lys Asp Thr Lys  
225 230 235 240  
Asn Ile Asp Val Val Arg Glu Leu Glu Ala Glu Phe Ser Ser Gly Ile  
245 250 255  
Gly Ser Gly Ser Met Asp Ser Ser Asp Ser Ser Asp Ser Ser Ser Ser  
260 265 270  
Ser Ser Asp Ser Ser Asp Ser Ser Asp Ser Ser Asp Ser Glu Ser Ser  
275 280 285  
Asp Asp Ser Glu Gly Gly Asp Asn Lys Val Arg Arg Ile Arg Arg His  
290 295 300  
Gln Tyr His Arg Arg Gln Leu Ser Tyr Ser Asp Asp Val Asn Gly Gly  
305 310 315 320  
Gly Arg Asn Ser Glu Lys Met Glu Met Asp Arg Val Thr His Ile Lys  
325 330 335  
Thr Glu His Ile Lys Arg Glu Asp Glu Pro Arg Tyr Glu Glu Arg Glu  
340 345 350  
Arg Tyr Ile His Pro Arg Arg Met Gln Val Pro Lys Asp Tyr Tyr Cys  
355 360 365  
Glu Gln Tyr Glu His Tyr Asp Ala Pro Ala Ala His His His Arg  
370 375 380  
His His Gln His Arg His Gln His Gln Arg His Phe Asn Gln Pro Arg  
385 390 395 400  
Ser Asn Asn Ser Ser Asp Val Thr Ala Tyr Val Asn Glu Asn Ser Pro  
405 410 415  
Trp Cys Arg Asp Arg Asn Ser Arg Phe Ser Pro Asn Asn Gly Gly Tyr  
420 425 430  
Asn Arg Ile Asn Ser Arg Tyr Thr Thr Phe Asp Pro Tyr Arg Tyr Gly  
435 440 445  
Ala Arg Arg Gly Arg Gly Gly Val Tyr  
450 455

<210> 186  
<211> 1014  
<212> DNA  
<213> SHRIMP

<400> 186  
atgtcctcct ctcaagggtt gaataataat atgtgcacca cagaaatcct gctgccccaa 60  
tgcacatcct cttccttgtc ttttagaggag agtgtggatt atttagaaaa ggattttgaa 120  
gaacttggaa tacctcttgt tgaaggaaag gaagtactac tggaaattgc ctacaaaaata 180  
ttaaacaaaa gggacacaat acgtgtaatt ggtgacgagc aaggagacgt atgtagcgtc 240  
ttctttcttc gttttggaaa gaagaagact tttaatccac aaacaaaaat gtggctagt 300  
aaactggcca atgctatcgc cctatccatg ggtgttgc ccagaacctgc ctgcacgtgt 360  
tccagaatga tgacgactgc aaagaagatc cctgttccag aatcatacaa aaatgttaat 420  
cgcaatatcc aaaaatttga agatgtacat tatatagata tcaattttca gtccttttga 480  
agagaacaga taggttttaag tgtattaggt aaaaatgatg tccaaaagaa gaagaaggaa 540  
gaaacccctt tctttgcacc ctttaataaa tctaaaatag gaggtgaatg catagaagat 600  
ttaaagtatg attctgagtc tgtttctatt ataagagatg tgttttaatt attgggtgaa 660  
atgcctactg aggatgtaaa gacatcaaga agttgtataa acccttccca caatgatacg 720  
aatcctagta tgagggttagt gtttcgtccc atgtactgga gaaattctaa gctgggtcatg 780  
gataaattat ccaaggaaaca agactcggct ttgattgaaa agtataatgg aggagaacat 840  
caacattgta catttgagg gagaaatgta ttattgtatt gtataactgc actatgtttt 900  
agctctgatt gtggatttaa aaagatgtta actaatgatg aaataaaaca attgatatgg 960  
tatttggtac ttttattttt tcatataatc tgtcctatta tacaatccaa atga 1014

WO 01/38351

272

PCT/US00/28888

<210> 187  
<211> 335  
<212> PRT  
<213> SHRIMP

<400> 187

Met	Ser	Ser	Ser	Gln	Gln	Asn	Asn	Met	Cys	Thr	Thr	Glu	Ile	Leu	Leu
1				5					10					15	
Pro	Lys	Cys	Thr	Ser	Ser	Ser	Leu	Ser	Leu	Glu	Glu	Ser	Val	Asp	Tyr
			20					25					30		
Leu	Glu	Lys	Asp	Phe	Glu	Glu	Leu	Gly	Ile	Pro	Leu	Val	Glu	Gly	Lys
		35					40					45			
Glu	Val	Leu	Leu	Glu	Phe	Ala	Tyr	Lys	Ile	Leu	Asn	Lys	Arg	Asp	Thr
	50					55					60				
Ile	Arg	Val	Ile	Gly	Asp	Glu	Gln	Gly	Asp	Val	Cys	Ser	Val	Phe	Phe
65				70					75						80
Leu	Arg	Phe	Gly	Lys	Lys	Lys	Thr	Phe	Asn	Pro	Gln	Thr	Lys	Met	Trp
			85					90					95		
Leu	Val	Lys	Leu	Ala	Asn	Ala	Ile	Ala	Leu	Ser	Met	Gly	Val	Val	Pro
			100					105					110		
Glu	Pro	Ala	Cys	Thr	Cys	Ser	Arg	Met	Met	Thr	Thr	Ala	Lys	Lys	Ile
		115					120					125			
Pro	Val	Pro	Glu	Ser	Tyr	Lys	Asn	Val	Asn	Arg	Asn	Ile	Gln	Lys	Phe
	130					135					140				
Glu	Asp	Val	His	Tyr	Ile	Asp	Ile	Asn	Phe	Gln	Ser	Phe	Val	Arg	Glu
145					150					155					160
Gln	Ile	Gly	Leu	Ser	Val	Leu	Gly	Lys	Asn	Asp	Val	Gln	Lys	Lys	Lys
			165					170					175		
Lys	Glu	Glu	Thr	Pro	Phe	Phe	Ala	Pro	Phe	Asn	Lys	Ser	Lys	Ile	Gly
			180					185					190		
Gly	Glu	Cys	Ile	Glu	Asp	Leu	Lys	Tyr	Asp	Ser	Glu	Ser	Val	Ser	Ile
		195					200					205			
Ile	Arg	Asp	Val	Phe	Asn	Leu	Leu	Gly	Glu	Met	Pro	Thr	Glu	Asp	Val
	210					215					220				
Lys	Thr	Ser	Arg	Ser	Cys	Ile	Asn	Pro	Ser	His	Asn	Asp	Thr	Asn	Pro
225					230					235					240
Ser	Met	Arg	Leu	Val	Phe	Arg	Pro	Met	Tyr	Trp	Arg	Asn	Ser	Lys	Leu
			245					250					255		
Val	Met	Asp	Lys	Leu	Ser	Lys	Glu	Gln	Asp	Ser	Ala	Leu	Ile	Glu	Lys
			260					265					270		
Tyr	Met	Gly	Gly	Glu	His	Gln	His	Cys	Ile	Ile	Gly	Gly	Arg	Asn	Val
		275					280					285			
Leu	Leu	Tyr	Cys	Ile	Thr	Ala	Leu	Cys	Phe	Ser	Ser	Asp	Cys	Gly	Phe
		290				295					300				
Lys	Lys	Met	Leu	Thr	Asn	Asp	Glu	Ile	Lys	Gln	Leu	Ile	Trp	Tyr	Leu
305					310					315					320
Val	Leu	Leu	Phe	Phe	His	Ile	Ile	Cys	Pro	Ile	Ile	Gln	Ser	Lys	
			325					330						335	

<210> 188  
<211> 3627  
<212> DNA  
<213> SHRIMP

<400> 188

atgggtgtc cagaagccaa aaaagtgtac gaaaatgcgt acggtgcaca gaatgggaga 60  
gtcatcaagg agaaaactgg atatgaagat tgttacgatg acgaagacga cgaggattat 120  
tgttcaggag aagaagattg cacaactagc tcacttctca aagccacttc tcttgccaac 180  
atcaactcca agaacttcct ggattttgga agaggcaaga aatcttcctc ttcttcacct 240



WO 01/38351

274

PCT/US00/28888

<213> SHRIMP

<400> 189

Met	Gly	Val	Pro	Glu	Ala	Lys	Lys	Val	Tyr	Glu	Asn	Ala	Tyr	Gly	Ala
1				5				10						15	
Gln	Asn	Gly	Arg	Val	Ile	Lys	Glu	Lys	Thr	Gly	Tyr	Glu	Asp	Cys	Tyr
			20				25						30		
Asp	Asp	Glu	Asp	Asp	Glu	Asp	Tyr	Cys	Ser	Gly	Glu	Glu	Asp	Cys	Thr
		35					40					45			
Thr	Ser	Ser	Leu	Leu	Lys	Ala	Thr	Ser	Leu	Ala	Asn	Ile	Asn	Ser	Lys
	50				55						60				
Asn	Phe	Leu	Asp	Phe	Gly	Arg	Gly	Lys	Lys	Ser	Ser	Ser	Ser	Ser	Pro
65				70						75					80
Thr	Cys	Asp	Tyr	Thr	Leu	Asp	Met	Val	Asp	Leu	Pro	Thr	Tyr	Asn	Val
				85				90						95	
Ser	Asp	Leu	Val	Met	Leu	Gly	Arg	Gln	Ile	Ala	Thr	Thr	Met	Leu	Lys
			100					105					110		
Gly	Gln	Lys	Asn	Met	Gly	Gln	Met	Ile	Leu	Phe	Ile	Asn	Thr	Thr	Asn
		115					120					125			
Gln	Gln	Ile	Ile	Asp	Val	Leu	His	Asp	Gly	Phe	Asn	Val	Ile	Arg	Glu
	130					135						140			
Glu	Asp	Thr	Met	His	Ser	Arg	Met	Gln	Asn	Lys	Lys	His	Ile	Tyr	Glu
145				150						155					160
Asn	Phe	Tyr	Cys	Arg	Asp	Glu	Lys	Lys	Val	Ile	Ser	Glu	Phe	Phe	Ser
				165					170					175	
Arg	Lys	Tyr	Lys	His	Glu	Lys	Ile	Lys	Ala	Arg	Ile	Glu	Arg	Val	Pro
			180					185					190		
Ile	Ile	Ile	Pro	Ser	Ser	Gln	Glu	Glu	Val	Asp	Trp	Leu	Thr	Glu	Pro
	195						200					205			
Pro	Ile	Glu	Asp	Met	Met	Met	Ala	Pro	Pro	Val	Ser	Asn	His	Lys	Met
	210					215					220				
Asp	Asp	Tyr	Glu	Gly	Leu	Asp	Tyr	Trp	Ile	Asn	Lys	His	Thr	Asp	Val
225					230					235					240
Met	Lys	Lys	Arg	Lys	Phe	Leu	Thr	Asn	Ser	Phe	Leu	Phe	Arg	Asn	Val
				245				250						255	
Pro	Thr	Thr	Ser	Phe	Asn	Ser	Ser	Pro	Thr	Ala	Val	Leu	Lys	Ser	Arg
			260					265					270		
Phe	Lys	Asp	Ala	Phe	Phe	Ala	Ser	Gln	Met	Glu	Gly	Val	Ile	Leu	Tyr
		275				280						285			
Tyr	Ala	Phe	Arg	Met	Ile	Arg	Val	Met	Lys	Asn	Leu	Leu	Lys	Ser	Lys
	290					295					300				
Asn	Leu	Lys	Gly	Arg	Tyr	Thr	Val	Leu	Phe	Thr	Asp	Gly	Lys	Ala	Pro
305					310					315					320
Ala	Ile	Lys	Met	Met	Thr	Arg	Ala	Lys	Arg	Gln	Ile	Arg	Gln	Glu	Arg
				325					330					335	
Ser	Lys	Glu	Lys	Ala	Lys	Ser	Arg	Asn	Glu	Asn	Cys	Leu	Asn	Arg	Lys
			340					345					350		
Thr	Asn	Asp	Leu	Leu	Phe	Tyr	Ser	Cys	Glu	Arg	Met	Met	Met	Arg	Leu
		355					360					365			
Pro	Gln	Gly	Leu	Met	Ala	Ser	Ala	Leu	Leu	Asp	Ile	Met	Arg	Ile	Pro
					375						380				
Val	Leu	Lys	Thr	Thr	Gly	Ser	Lys	Cys	Met	Tyr	Leu	Ser	Asn	Ala	Ser
385					390					395					400
Phe	Thr	Glu	Ala	Glu	Asp	Asp	Ile	Val	Arg	Leu	Thr	Ser	Cys	Leu	Leu
				405					410					415	
Asn	Leu	Glu	Thr	Pro	Gly	Lys	His	Phe	Ser	Leu	Leu	Glu	Lys	Arg	Lys
			420					425					430		
Ile	Tyr	Asp	Ser	Tyr	Asn	Met	Ser	Gly	Asn	Arg	Lys	Glu	Ser	Lys	Arg
		435					440					445			
Trp	Glu	Asp	Leu	Leu	Asn	Val	Leu	Lys	Gln	His	Thr	Asn	Asp	Glu	Asn
	450					455						460			

Gln Thr Leu Ser Met Asn Leu Phe Ser His Asp Ser Asp Val Leu Val  
 465 470 475 480  
 Lys Trp Asn Leu Met Val Gly His His Lys Asn Val Cys Arg Leu Thr  
 485 490 495  
 Gly Thr Gln Phe Lys Asp Ser Glu Thr Phe Leu Lys Ile Gly His Val  
 500 505 510  
 Lys Phe Phe Arg Cys Met Asn Ser Asn Ser Ser Gly Glu Asn Gln Ala  
 515 520 525  
 Asn Glu Leu Gly Gly Phe Ala Ala Lys Arg Arg Thr Lys Pro Asn Thr  
 530 535 540  
 Ile Tyr Asn Leu Ala Glu Ser Pro Leu Met Leu Ser Pro Glu Ser Thr  
 545 550 555 560  
 Leu Leu Ile Met Leu Thr Lys Gly Ser Asp Tyr Asn Ser Ala Ile Val  
 565 570 575  
 Ser Asn Cys Glu Tyr Asp Thr Trp Val Arg Lys Glu Val Ala Val Phe  
 580 585 590  
 Glu Asn Thr Tyr Cys Thr Cys Val Gly Gly Trp Glu Ile Phe Leu Ser  
 595 600 605  
 Glu Gln Glu Ala Arg Lys Asn Asn Lys Asp Cys Asp Asp Ser Val Gly  
 610 615 620  
 Asn Ile Ser Met Gly Asn Leu Ser Lys Ser Asn Cys Arg Lys Cys Asp  
 625 630 635 640  
 Lys Lys Leu Val Leu Pro Phe Trp Thr Ile Lys Phe Phe Tyr Leu Ser  
 645 650 655  
 Gln Ala Ile Asp Phe Val Arg Asp Pro Leu Gln Leu Cys Phe Pro Pro  
 660 665 670  
 Thr His Leu Ile Asp Leu Glu Thr Asp Val Ser Leu Lys His Ala Leu  
 675 680 685  
 His Arg Ala Val Asn Ala Ala Ala Asn Val Met Ser Tyr Leu Thr Met  
 690 695 700  
 Gly Ser Phe Asn Gln Arg Val Phe Gly Thr Ile Thr Thr Leu Ser Asp  
 705 710 715 720  
 Ile Ser Ile His Leu Ser Gly Ala Asn Asn Asn Glu Ser Lys Asn Thr  
 725 730 735  
 Gly Ser Asp Val Glu Ser Asp Thr Glu Asp Leu Ile Pro Phe Ser Asn  
 740 745 750  
 Asn Lys Arg Lys Ser Gly Asn Asp Pro Gln Lys Ser Thr Arg Lys Lys  
 755 760 765  
 Ser Lys Val Asn Ala Thr Arg Lys Ser Ala Pro Val Thr Lys Lys Leu  
 770 775 780  
 Ser Ser Ser Val Phe Glu Ser Ile Arg Gly Phe Phe Glu Ser His Thr  
 785 790 795 800  
 Glu Gly Gly Ile Ile Asn Asp Arg Gly Ile Leu Thr Lys Glu Arg Ile  
 805 810 815  
 Asp Val Phe Gly Asn Asn Leu Asp Thr Asn Pro Glu Ala Leu Gly Glu  
 820 825 830  
 Glu Asn Gly Gly Gly Gly Ile Val Ser Ser Ile Pro Gly Leu Ser  
 835 840 845  
 Thr Glu Gln Thr Ser Ile Leu Lys Thr Glu Gln Asn Asn Ser Thr Ser  
 850 855 860  
 Asp Phe Leu Asp Phe Phe Lys Lys Phe Asn Glu Met Asp Asp Val Glu  
 865 870 875 880  
 Glu Glu Glu Glu Lys Met Glu Glu Gly Glu Lys Glu Glu Glu Glu Ala  
 885 890 895  
 Asp Leu Glu Thr Asp Asp Trp Leu Asp Glu Ala Arg Lys Ala Phe Glu  
 900 905 910  
 Tyr Lys Asp Ser Asp Phe Leu Glu Ala Val Thr Ala Thr Asn Glu  
 915 920 925  
 Met Thr Ser Ser Leu Ala Lys Asn Asn Ile Glu Glu Asp Glu His Ser  
 930 935 940  
 Arg Cys Ser Val Ser Ser Lys Leu Asn Asn Lys Gln Pro Val Met Asp

```

945          950          955          960
Glu Glu Lys Trp Ala Glu Ile Val Asn Glu Phe Asp Lys Cys Ile Ser
          965          970          975
Leu Asp Asn Ile Thr Tyr Asn Asp Asn Ser Leu Leu Ser Arg Leu Ser
          980          985          990
Gly Val Leu Met Asp Ala Asn Lys Arg Glu Asp Gly Asn Asn Ser Asn
          995          1000          1005
Val Val Leu Tyr Glu Pro Val Gln Gly Ile Asp Asp Glu Arg Phe Ser
          1010          1015          1020
Gly Val Pro Tyr Ser Val Lys Thr Met Asn Leu Leu Val Ile Val Tyr
1025          1030          1035          1040
Met Asn Met Cys Gly Leu Glu Asp Asn Thr Ile Val Tyr Gln Gln Leu
          1045          1050          1055
Met Pro Ile Ile His Ser Glu Phe Cys Gly Lys Thr Glu Glu Asp Lys
          1060          1065          1070
Ile Cys Thr Asp Arg Thr Asn Phe Met Ser Ala Ala Leu Glu Tyr Thr
          1075          1080          1085
Met Leu Gln Tyr Met Pro Glu Leu Lys Lys Thr Pro Arg Ile Lys Gln
          1090          1095          1100
Ile Lys Arg Lys Asn Trp Glu Arg Ile Pro Lys Val Leu Asp Asp Phe
1105          1110          1115          1120
Lys Asp Lys Val Ser Thr Cys Thr Asp Asn Tyr Asn Lys Leu Leu Ala
          1125          1130          1135
Thr Leu Asn Lys Glu Gly Lys Ile Pro Ser Glu Asn Thr Lys Trp Leu
          1140          1145          1150
Pro Ser Gln Gly Gln Phe Met Pro Val Leu Gly Val Ala Ile Ser Lys
          1155          1160          1165
Pro Trp Ser Pro Leu Thr Leu Trp Ser Ser Phe Tyr Leu Gln His Gln
          1170          1175          1180
Gln Arg Gln Asp Val Ser Leu Thr Asn Ile Thr Pro Pro Asn Ser Pro
1185          1190          1195          1200
Arg Pro Glu Gln

```

<210> 190  
<211> 414  
<212> DNA  
<213> SHRIMP

```

<400> 190
atggatggag attcttcctc cttacagtta ttgagtgaat ccgaatttga ttacgtggtc 60
gagacactta aggaacaagg tgtgtgggaa ttggccctag aagtgttcaa cgaggtgtcc 120
aattctattg agactgtgaa agaggaagaa gactacactg ttcttcgatc tagaaactac 180
ttccctactg aatctataac actctacaaa caacaacagg aagaggaaga aagtaccct 240
attaagaaga ggaaactcgc ttctggcaag tctccgagaa gtctctgtag agagctgcgt 300
ttgctgcaga ttccaagcac tacaaccttt aaagcagctc cacgaagttc ttctaggagg 360
ggtaaaaaca ccagactacg cagagtgtgt aaaaattacg gcgcccatca gtga 414

```

<210> 191  
<211> 137  
<212> PRT  
<213> SHRIMP

```

<400> 191
Met Asp Gly Asp Ser Ser Ser Leu Gln Leu Leu Ser Glu Ser Glu Phe
1      5      10      15
Asp Tyr Val Val Glu Thr Leu Lys Glu Gln Gly Val Trp Glu Leu Ala
20     25     30
Leu Glu Val Phe Asn Glu Val Ser Asn Ser Ile Glu Thr Val Lys Glu
35     40     45

```



WO 01/38351

PCT/US00/28888

277

Glu	Glu	Asp	Tyr	Thr	Val	Leu	Arg	Ser	Arg	Asn	Tyr	Phe	Pro	Thr	Glu
	50					55				60					
Ser	Ile	Thr	Leu	Tyr	Lys	Gln	Gln	Gln	Glu	Glu	Glu	Glu	Ser	Thr	Pro
65					70				75						80
Ile	Lys	Lys	Arg	Lys	Leu	Ala	Ser	Gly	Lys	Ser	Pro	Arg	Ser	Leu	Cys
				85				90						95	
Arg	Glu	Leu	Arg	Leu	Leu	Gln	Ile	Pro	Ser	Thr	Thr	Thr	Phe	Lys	Ala
			100					105					110		
Ala	Pro	Arg	Ser	Ser	Ser	Arg	Arg	Gly	Lys	Asn	Thr	Arg	Leu	Arg	Arg
		115					120					125			
Val	Cys	Lys	Asn	Tyr	Gly	Ala	His	Gln							
	130					135									

<210> 192  
 <211> 924  
 <212> DNA  
 <213> SHRIMP

<400> 192  
 atgtggtggt ctacacacct ttcatatagt gagtttttta ctccctcttca atctcaaaaa 60  
 cttttcagaa attttttcag ggcactcgag tttagggggt ggaccgctag ctcgaccgaa 120  
 tgtcaagttc cgaggggtgga cctctgggtc gggccaatgt cagattacac cagaaattgt 180  
 tggttccaga aacgtacatt aactttttgtg tgtttctgga acaggcggtt ctggagactg 240  
 gtcgacccag aaatgagagg gtataacctt ctgtttttcac tagagaattt tactcttctt 300  
 ctatctcaaa aactttttcaa aaattttttc agggcactcc agtttagggg gtggaccgct 360  
 agctcgaccg aatgtcaagt tccgaggggt gacgctggg tcgggccaat gtcagattac 420  
 accagaaatg taatagctcc agaaacgtac attaaacttt gtgtgtttct ggaacaggcg 480  
 tttctggaga ctggtcgacc cagaaatgag aggtataacc cttctgtttt cactagagaa 540  
 ttttactctt cctccatctc aaaaactttt caaaaatttt ttagggcgct ccagtttagg 600  
 ggggtggaccg cttagctcgac cgaatgtcaa gttccgaggg tggacctctg ggtcgggcca 660  
 atgtcagatt acaccagaaa tgtaatagct ccagaaatag aggaagtatc ttatggccat 720  
 ttctggacca ggtgtttctg gacgaagata ttgcttgatg gcaaccctct cccctccct 780  
 ccccttttta aaaagggccc acgtgtatat aatgactgta ccacacctca ttcaaaccat 840  
 cacaaccacc accaccacca tggaagaaca tctatccttc aacaaaccct ctccagaaaa 900  
 tggagtagtc ttctttgact ttag 924

<210> 193  
 <211> 305  
 <212> PRT  
 <213> SHRIMP

<400> 193  
 Met Trp Cys Ser Thr His Leu Ser Tyr Ser Glu Phe Phe Thr Pro Leu  
 1 5 10 15  
 Gln Ser Gln Lys Leu Phe Arg Asn Phe Phe Arg Ala Leu Glu Phe Arg  
 20 25 30  
 Gly Trp Thr Ala Ser Ser Thr Glu Cys Gln Val Pro Arg Val Asp Leu  
 35 40 45  
 Trp Val Gly Pro Met Ser Asp Tyr Thr Arg Asn Cys Trp Phe Gln Lys  
 50 55 60  
 Arg Thr Leu Thr Phe Val Cys Phe Trp Asn Arg Arg Phe Trp Arg Leu  
 65 70 75 80  
 Val Asp Pro Glu Met Arg Gly Tyr Asn Leu Leu Phe Ser Leu Glu Asn  
 85 90 95  
 Phe Thr Leu Pro Leu Ser Gln Lys Leu Phe Lys Asn Phe Phe Arg Ala  
 100 105 110  
 Leu Gln Phe Arg Gly Trp Thr Ala Ser Ser Thr Glu Cys Gln Val Pro  
 115 120 125  
 Arg Val Asp Arg Trp Val Gly Pro Met Ser Asp Tyr Thr Arg Asn Val  
 130 135 140

WO 01/38351

PCT/US00/28888

278

Ile Ala Pro Glu Thr Tyr Ile Asn Phe Cys Val Phe Leu Glu Gln Ala  
 145 150 155 160  
 Phe Leu Glu Thr Gly Arg Pro Arg Asn Glu Arg Val Tyr Pro Ser Val  
 165 170 175  
 Phe Thr Arg Glu Phe Tyr Ser Ser Ser Ile Ser Lys Thr Phe Gln Lys  
 180 185 190  
 Phe Phe Arg Ala Leu Gln Phe Arg Gly Trp Thr Ala Ser Ser Thr Glu  
 195 200 205  
 Cys Gln Val Pro Arg Val Asp Leu Trp Val Gly Pro Met Ser Asp Tyr  
 210 215 220  
 Thr Arg Asn Val Ile Ala Pro Glu Ile Glu Glu Val Ser Tyr Gly His  
 225 230 235 240  
 Phe Trp Thr Arg Cys Phe Trp Thr Lys Ile Leu Leu Asp Gly Asn Pro  
 245 250 255  
 Leu Pro Leu Pro Pro Phe Lys Lys Gly Pro Arg Val Tyr Asn Asp  
 260 265 270  
 Cys Thr Thr Pro His Ser Asn His His Asn His His His His His Gly  
 275 280 285  
 Arg Thr Ser Ile Leu Gln Gln Thr Leu Ser Arg Lys Trp Ser Ser Leu  
 290 295 300  
 Leu  
 305

<210> 194  
 <211> 447  
 <212> DNA  
 <213> SHRIMP

<400> 194  
 atgacccacc ttgtccttct tatectatcc ttatccttgt cccctgtcta ccaccacctc 60  
 accccatatt tctcaccoca tctcacctat acccccatct caccatcac ctctatatcc 120  
 cccacacctta ttactcgtct ccagtttcaa caccctgttc ttgccgagcc aacccataac 180  
 cagatctgga cccagtcctt cctttttatc cctaaccggc accatttatg cccccaggcg 240  
 ctacggtgt atataaggcg gcgcggccag gccagaagca tcagttctct gcaagccagc 300  
 agaagagcaa cacaacaagc actctctctc cttctaccta gaagagacct gccaatatct 360  
 aagctacaag aatggcctct ccagcccccg ccgcaccaag tccttacacc atgttggtact 420  
 ctaagttact tagttctgag gaactaa 447

<210> 195  
 <211> 146  
 <212> PRT  
 <213> SHRIMP

<400> 195  
 Met Thr His Leu Val Leu Leu Ile Leu Ser Leu Ser Leu Ser Pro Val  
 1 5 10 15  
 Tyr His His Leu Thr Pro Tyr Leu Ser Pro His Leu Thr Tyr Thr Pro  
 20 25 30  
 Ile Ser Pro Ile Thr Ser Ile Phe Pro His Leu Ile His Ser Leu Gln  
 35 40 45  
 Phe Gln His Pro Val Leu Ala Glu Pro Thr His Asn Gln Ile Trp Thr  
 50 55 60  
 Pro Val Phe Pro Phe Ile Pro Asn Arg His His Leu Cys Pro Gln Ala  
 65 70 75 80  
 Val Tyr Ile Arg Arg Arg Gly Gln Ala Arg Ser Ile Ser Ser Leu Gln  
 85 90 95  
 Ala Ser Arg Arg Ala Thr Gln Gln Ala Leu Ser Leu Leu Leu Pro Arg  
 100 105 110  
 Arg Asp Leu Pro Ile Leu Lys Leu Gln Glu Trp Pro Leu Gln Pro Pro  
 115 120 125

WO 01/38351

279

PCT/US00/28888

Pro His Gln Val Leu Thr Pro Cys Trp Thr Leu Ser Tyr Leu Val Leu  
 130 135 140  
 Arg Asn  
 145

<210> 196  
 <211> 339  
 <212> DNA  
 <213> SHRIMP

<400> 196  
 atgttggaact ctaagttact tagttctgag gaactaaagg aactaacttc atacgtctcg 60  
 actagctctc gccggtctga tatgaagaaa cacttgctcc atctattcga ggagcagag 120  
 aagatcttcc aattcatata aggttaagcac aagttctcac tatacacttt ggactttgaa 180  
 attttctatg ttatgctgaa tattttgttg gttgaagtga aaaatattct aagtccaatt 240  
 cctttactct ttgacagaaa tctccaacca gtacggagac tatggatggt tcacaatggc 300  
 cccgcctcac ctgaacgctg cagccgatct cttggataa 339

<210> 197  
 <211> 110  
 <212> PRT  
 <213> SHRIMP

<400> 197  
 Met Leu Asp Ser Lys Leu Leu Ser Ser Glu Glu Leu Lys Glu Leu Thr  
 1 5 10 15  
 Ser Tyr Val Ser Thr Ser Ser Arg Arg Ser Asp Met Lys Lys His Leu  
 20 25 30  
 Leu His Leu Phe Glu Glu His Glu Lys Ile Phe Gln Phe Ile Gln Gly  
 35 40 45  
 Lys His Lys Phe Ser Leu Tyr Thr Leu Asp Phe Glu Ile Phe Tyr Val  
 50 55 60  
 Met Leu Asn Ile Leu Leu Val Glu Val Lys Asn Ile Leu Ser Pro Ile  
 65 70 75 80  
 Pro Leu Leu Phe Asp Arg Asn Leu Gln Pro Val Arg Arg Leu Trp Met  
 85 90 95  
 Phe His Asn Gly Pro Asp Glu Arg Cys Ser Arg Ser Leu Gly  
 100 105 110

<210> 198  
 <211> 3438  
 <212> DNA  
 <213> SHRIMP

<400> 198  
 atggccccgc ctcacctgaa cgctgcagcc gatctcttgg ataagggtgat gtccggacact 60  
 ctatctcccc aaggcgccca aacttcgctc ccggtctgtt gtgttggtgc gaaggttgtg 120  
 aaggcactgg tgagcttttg ccagaagacc cgcttcacca ccaacattgt gatgagagaa 180  
 gttaaagcca tggagtcca aggagacgat tttaactact ctgccttgtg tgcaagtatg 240  
 ccccaacgcc ccgtgactga gaggcagatg ttcgccctta tgaagagtga ggacgaagaa 300  
 atgggagtgt ctgcaaactt ctctccagtc tctgatgacg tcatcaaccc ttcaagcctc 360  
 cctctggac aagaagtcca ctcatcaact tccgctcaaa tttctggtat gtttcaaaac 420  
 gtgtggagtt tgcttgaaga gtgtggtagt ggctctaata gtaatagttc ccctgtctct 480  
 aggacagtct tagtttgac cctgtttata atccaagtgt tcaagttttt ggtgactaaa 540  
 gtgtctaata tgaacgtact taaccagttg tttggacatg ttgttttttg atcacttgat 600  
 gtggctccaa gtaataataa tagtgtccca tcaactgttg ttaacaacaa caacaaaccc 660  
 tcgacctcta ataataagtaa caacatcagt aacaagcgtg ttggtggtag taataacagt 720  
 ggcggcggaa gatcaaagaa agttacagcc acagccaaaa atccctttta taatgtagat 780  
 ggggacaatc atggcatgtt tgccggtgcc cctgttgatg ttaatttgga tgactttgtt 840

WO 01/38351

280

PCT/US00/28888

```

tttccccaag ttgaaactct tacaagtaag agcaccatcc ctaaagaaga ggtaaatgta 900
gatgaagatt tgagtaaaat gtgccgtaaa actgccctta cccccctaga aattcatacc 960
tttaatgtgt tcatctctga gattaacccc tccaaatatg accgttcaat gttttgcaag 1020
ggatttttga ctgcatggga taagtttgta gaggggggata ctgctggcgt taaacgcttc 1080
cgtaactata tcttcactcg ctcaaactat gcctcagccg ccagggccgt gtatgaagcg 1140
tcaattaaag ggactgttta ttataatgac aagtcaaaagt ttctgttcca cgataatggt 1200
aaccctgatac tggacaagag ctggggtaac aagaatggga agaaacctag actcccagct 1260
aacttgatgg cattcatggg tattgacatt gtaaagggtg gcgctaaggg gattcaaaaag 1320
tatatgtttg caaagcaatt ccaacatccg gaagtgggaag aacttggtgc tcctatggct 1380
gtatacgcaa aggttgccgc aggattgaag tcggggactt tgtttgatga ctgggacctg 1440
cctgaatacg aaaattgtca gtttatcaag tatgacacag aagggtgcaa aaagcacagt 1500
gagttatacg ccaacaact tctccgcaca ggacttaatc aatacaataa actggaagag 1560
ggacagagtg cattcccatt tgcaaatatt gtgacggtaa catccgcctc tagtgatgat 1620
attcacggtg acacaatcat tgaattgatg tacaagacaa aggatggcgt aaagggagtc 1680
tcaaaaattg aggacgaaaa catcatcaag gtgaatccag cagaagaaaa gaagaataat 1740
agagtacaag ccgagaagac cctgtatttt gagattgatt ccgatgatga ggtgtgtgag 1800
agaacagagg aagaattctt caggcctaca tctgtgtgtg ctgccccgac aacacccttc 1860
gtaccttcta atgtggagga agaggaagag gaagaagagc agatggaaga agaggaggaa 1920
gaggaagtag aaaggggaaga aggatctgat aaggaagatg acggagacgc accagcacag 1980
gaagaaatgg aggaggagaa ggaagaagaa caacaacaac agccagaaga agaaagcaat 2040
ggtaatgaga accaagaaga agaacaacaa caacaacaac aaccagaaag agaagaggag 2100
aataaggatg cagatagtag cagcgacagt gatagcagca gcagcagtag tagcagcagt 2160
agcagcagta gtagtagtag cagcagtagt agcagcagca gtagtagcag tgaaaatgaa 2220
gctgaaaaga agaaagaaga ggaagtacct gccaaagattc agaagagaaa gaggctaagt 2280
gaaaggccat cagaagctgc ttcctctccc aagagaatga gtagtagaaga agaacaacaa 2340
caacaactat caccatcatt ggacatactc cagactgcag ttgatgagat gatggaagaa 2400
attcctgcgc ctgagcctat cgttgctaca acctcaccca aggcagcgac acttgctactc 2460
aagacaggat ttagttactc ttcattctga agaggagatg acctttcagt agctggtaat 2520
acttccccta ctgaaccagc agctgtgccc gctgctgcca cttgcacttc cgatgttgga 2580
aatgactttt tggacatggt ggacggttta cctggcgata tagtaatgca acctggcgaa 2640
tgcgacgtga ccgcaaaatt ctttgagggc atcacccctac cagatggtac tgataatgaa 2700
tgcacagggt tcgatgatct tcttaaagcc acccgagactg ataacattat aaccaccaca 2760
tgctttacct ccccgattca ccttctagc aactcagccc ccagaaagga tattgataat 2820
tgcagttcta ttaagaggtc tagggcaggt tcactttttg acactgatga tgatagttaa 2880
acaaatgagg ttgaaaagga agcccctaaa cgtaagaagc acttgaaaaa gaggcgtaac 2940
aagtcccacc gtggttctctc tggttctgtc tcttcttctc attgtatgag tagtgatgaa 3000
gaatcagagg atgaaaggga tatgaaatca acatcaaagg ttcacaagtc accaaaagct 3060
catgtttaaac attcccctaa atatgatgct gtaaatagtg atgtaataaa ctcatacaac 3120
aatgttaata gtacaacatg catgtcgtca tcagatagtg atgcagaagc acagcctaaa 3180
agccataata aaagccactc tcgtaaacac tcttcttctc ccacaagtga taagaaacag 3240
aaccaacaat gtcfaatcaa tactcaaaat gtcaagaaga ctgttgatga gtctccacct 3300
agtttttagaa gtttttagtcc taagaaagat gagcttggtg atttcttgtc acgcaagcac 3360
acaaagccag ttaggcccta taacaagaag cgtgataatg ttaacaccac taataatgta 3420
gtacagaggt ctgcctga 3438

```

<210> 199  
 <211> 1141  
 <212> PRT  
 <213> SHRIMP

<400> 199  
 Met Ala Pro Pro His Leu Asn Ala Ala Asp Leu Leu Asp Lys Val  
 1 5 10 15  
 Met Ser Gly Pro Leu Ser Pro Glu Gly Ala Gln Thr Ser Ser Pro Ala  
 20 25 30  
 Ala Cys Val Gly Ala Lys Val Val Lys Ala Leu Val Ser Phe Cys Gln  
 35 40 45  
 Lys Thr Arg Phe Thr Thr Asn Ile Val Met Arg Glu Val Lys Ala Met  
 50 55 60  
 Glu Phe Gln Gly Asp Asp Phe Asn Tyr Ser Ala Leu Cys Ala Ser Met  
 65 70 75 80

Pro Gln Arg Pro Val Thr Glu Arg Gln Met Phe Ala Leu Met Lys Ser  
 85 90 95  
 Glu Asp Glu Glu Met Gly Val Ser Ala Asn Phe Ser Pro Val Ser Asp  
 100 105 110  
 Asp Val Ile Asn Pro Ser Ser Leu Pro Ser Gly Gln Glu Val Asp Ser  
 115 120 125  
 Ser Thr Ser Ala Gln Ile Ser Gly Met Phe Gln Asn Val Trp Ser Leu  
 130 135 140  
 Leu Glu Glu Cys Gly Ser Gly Ser Asn Ser Asn Ser Ser Pro Val Ser  
 145 150 155 160  
 Arg Thr Val Leu Val Cys Thr Leu Phe Ile Ile Gln Val Phe Lys Phe  
 165 170 175  
 Leu Val Thr Lys Val Ser Asn Val Asn Val Leu Asn Gln Leu Phe Gly  
 180 185 190  
 His Val Val Phe Gly Ser Leu Asp Val Ala Pro Ser Asn Asn Ser  
 195 200 205  
 Val Pro Ser Thr Val Val Asn Asn Asn Asn Lys Pro Ser Thr Ser Asn  
 210 215 220  
 Asn Ser Asn Asn Ile Ser Asn Lys Arg Val Gly Gly Ser Asn Asn Ser  
 225 230 235 240  
 Gly Gly Gly Arg Ser Lys Lys Val Thr Ala Thr Ala Lys Asn Pro Phe  
 245 250 255  
 Asn Asn Val Asp Gly Asp Asn His Gly Met Phe Ala Gly Ala Pro Val  
 260 265 270  
 Asp Val Asn Leu Asp Asp Phe Val Phe Pro Gln Val Glu Thr Leu Thr  
 275 280 285  
 Ser Lys Ser Thr Ile Pro Lys Glu Glu Val Asn Val Asp Glu Asp Leu  
 290 295 300  
 Ser Lys Met Cys Arg Lys Thr Ala Leu Thr Pro Leu Glu Ile His Thr  
 305 310 315 320  
 Phe Asn Val Phe Ile Ser Glu Ile Asn Pro Ser Lys Tyr Asp Arg Ser  
 325 330 335  
 Met Phe Cys Lys Gly Phe Leu Thr Ala Trp Asp Lys Phe Val Glu Gly  
 340 345 350  
 Asp Thr Ala Gly Val Lys Arg Phe Arg Asn Tyr Ile Leu Thr Arg Ser  
 355 360 365  
 Asn Tyr Ala Ser Ala Ala Arg Ala Val Tyr Glu Ala Ser Ile Lys Gly  
 370 375 380  
 Thr Val Tyr Tyr Asn Asp Lys Ser Lys Phe Leu Phe His Asp Asn Val  
 385 390 395 400  
 Asn Pro Asp Leu Asp Lys Ser Trp Gly Asn Lys Asn Gly Lys Lys Pro  
 405 410 415  
 Arg Leu Pro Ala Asn Leu Met Ala Phe Met Gly Ile Asp Ile Val Lys  
 420 425 430  
 Val Cys Ala Lys Gly Ile Gln Lys Tyr Met Phe Ala Lys Gln Phe Gln  
 435 440 445  
 His Pro Glu Val Glu Glu Leu Val Pro Pro Met Ala Val Tyr Ala Lys  
 450 455 460  
 Val Ala Ala Gly Leu Lys Ser Gly Thr Leu Phe Asp Asp Trp Asp Leu  
 465 470 475 480  
 Pro Glu Tyr Glu Asn Cys Gln Phe Ile Lys Tyr Asp Thr Glu Gly Cys  
 485 490 495  
 Lys Lys His Ser Glu Leu Tyr Ala Lys Gln Leu Leu Arg Thr Gln Gln  
 500 505 510  
 Tyr Asn Lys Leu Glu Glu Gly Gln Ser Ala Phe Pro Phe Ala Asn Ile  
 515 520 525  
 Val Thr Val Thr Ser Ala Ser Ser Asp Asp Ile His Gly Asp Thr Ile  
 530 535 540  
 Ile Glu Leu Met Tyr Lys Thr Lys Asp Gly Val Lys Gly Val Ser Lys  
 545 550 555 560  
 Ile Glu Asp Glu Asn Ile Ile Lys Val Asn Pro Ala Glu Glu Lys Lys

```

                    565                    570                    575
Asn Asn Arg Val Gln Ala Glu Lys Thr Leu Tyr Phe Glu Ile Asp Ser
580                    585                    590
Asp Asp Glu Val Cys Glu Arg Thr Glu Glu Glu Phe Phe Arg Pro Thr
595                    600                    605
Ser Val Val Ala Ala Pro Thr Thr Pro Leu Val Pro Ser Asn Val Glu
610                    615                    620
Glu Glu Glu Glu Glu Glu Glu Gln Met Glu Glu Glu Glu Glu Glu
625                    630                    635
Val Glu Arg Glu Glu Gly Ser Asp Lys Glu Asp Asp Gly Asp Ala Pro
645                    650                    655
Ala Gln Glu Glu Met Glu Glu Glu Lys Glu Glu Glu Gln Gln Gln Gln
660                    665                    670
Pro Glu Glu Glu Ser Asn Gly Asn Glu Asn Gln Glu Glu Glu Gln Gln
675                    680                    685
Gln Gln Gln Gln Pro Glu Arg Glu Glu Glu Asn Lys Asp Ala Asp Ser
690                    695                    700
Asp Ser Asp Ser Asp Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser
705                    710                    715                    720
Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser
725                    730                    735
Asn Glu Ala Glu Lys Lys Lys Glu Glu Glu Val Pro Ala Lys Ile Gln
740                    745                    750
Lys Arg Lys Arg Leu Ser Pro Ser Glu Ala Ala Ser Ser Pro Lys Arg
755                    760                    765
Met Arg Val Glu Glu Glu Gln Gln Gln Leu Ser Pro Ser Leu Asp
770                    775                    780
Ile Leu Gln Thr Ala Val Asp Glu Met Met Glu Glu Ile Pro Ala Pro
785                    790                    795                    800
Glu Pro Ile Val Ala Thr Thr Ser Pro Lys Ala Ala Thr Leu Ala Leu
805                    810                    815
Lys Thr Gly Phe Ser Tyr Ser Ser Phe Val Arg Gly Asp Asp Leu Ser
820                    825                    830
Val Ala Gly Asn Thr Ser Pro Thr Glu Pro Ala Ala Val Pro Ala Ala
835                    840                    845
Ala Thr Cys Thr Ser Asp Val Gly Asn Asp Phe Leu Asp Met Leu Asp
850                    855                    860
Gly Leu Pro Gly Asp Ile Val Met Gln Pro Gly Glu Cys Asp Val Thr
865                    870                    875                    880
Ala Lys Phe Phe Glu Gly Ile Thr Leu Pro Asp Gly Thr Asp Asn Glu
885                    890                    895
Cys Thr Gly Phe Asp Asp Leu Leu Lys Ala Thr Glu Thr Asp Asn Ile
900                    905                    910
Ile Thr Thr Thr Cys Phe Thr Ser Pro Ile His Pro Ser Ser Asn Ser
915                    920                    925
Ala Pro Arg Lys Asp Ile Asp Asn Cys Ser Ser Ile Lys Arg Ser Arg
930                    935                    940
Ala Gly Ser Leu Phe Asp Thr Asp Asp Asp Ser Glu Thr Asn Glu Val
945                    950                    955                    960
Glu Lys Glu Ala Pro Lys Arg Lys Lys His Leu Lys Lys Arg Arg Asn
965                    970                    975
Lys Ser His Arg Gly Ser Ser Gly Ser Ala Ser Ser Ser His Cys Met
980                    985                    990
Ser Ser Asp Glu Glu Ser Glu Asp Glu Arg Asp Met Lys Ser Thr Ser
995                    1000                    1005
Lys Val His Lys Ser Pro Lys Ala His Val Lys His Ser Pro Lys Tyr
1010                    1015                    1020
Asp Ala Val Asn Ser Asp Val Asn Asn Ser Tyr Asn Asn Val Asn Ser
1025                    1030                    1035                    1040
Thr Thr Cys Met Ser Ser Ser Asp Ser Asp Ala Glu Ala Gln Pro Lys
1045                    1050                    1055

```

Ser	His	Asn	Lys	Ser	His	Ser	Arg	Lys	His	Ser	Ser	Ser	Ser	Thr	Ser
1060						1065					1070				
Asp	Lys	Lys	Gln	Asn	Gln	Gln	Cys	Ser	Ile	Asn	Thr	Gln	Asn	Val	Lys
1075						1080					1085				
Lys	Thr	Val	Val	Gln	Ser	Pro	Pro	Ser	Phe	Arg	Ser	Phe	Ser	Pro	Lys
1090						1095					1100				
Lys	Asp	Glu	Leu	Gly	Asp	Phe	Leu	Ser	Arg	Lys	His	Thr	Lys	Pro	Val
1105						1110					1115				
Arg	Pro	Tyr	Asn	Lys	Lys	Arg	Asp	Asn	Val	Asn	Thr	Thr	Asn	Asn	Val
1125						1130					1135				
Val	Gln	Arg	Ser	Ala											
1140															

<210>	200
<211>	1875
<212>	DNA
<213>	SHRIMP

<400>	200						
atgtcctcga	cgacctctcc	atcgtctttg	tgggatgatg	acgatgatga	tgacgaagaa	60	
gacgaaaaag	atgtcaagca	agaagttctg	aaccgtcccc	ccattttttc	gtacatggaa	120	
actgatctct	ttagtgataa	cgatgaggac	gataacaagg	gagaagaaga	atgttttggg	180	
tcaaacctttg	atatgttttg	tgatttcagat	aacatgccat	caacttctac	tgcccccttc	240	
cctcctccct	ctacaacaac	accacttctt	actcctcgat	ccatcatgga	tactgattcg	300	
gatgaatgtg	acgaagaagg	agcagcagca	gcatacagcac	cgtctattgc	cgcctcttct	360	
tctatccctg	tcgggatctc	tgaagctgaa	ttgaaaaaaa	tggaaaaaga	aaagagggaag	420	
gaatttaaga	aactcaaaaa	gatgatgaaa	gattcctctc	ctcacctata	tgtaggagga	480	
gaacctcctg	tgcgagcaga	ttataaaaaca	agggcaaaaca	tttcccttta	taaagttgac	540	
cctagtatcg	atatgtgctg	tgtcgccctt	cctcaatttt	gcgctgaatt	gcccacccca	600	
tccatagatg	tgtatacttc	ttcctatgta	tttctctctc	ccacacctgc	catgcataat	660	
aagaaagggt	ccaagaaatg	tcaattcctt	aaggggagaa	aggctttgag	gaaatggatt	720	
cacgagaatg	tatgcattgc	ccctcccggt	aaaaggggag	gtgtattttt	ggctcatttg	780	
gaacaaagat	tcttggtctg	acatggagat	gaatacaagg	tccaagggat	gtttgtttca	840	
agagtattga	acaaagcttt	ccccaatctg	attgctcgtg	cagacacact	gtgcagtgat	900	
atgacattct	atactaacct	ttgttggtata	gttaatggag	ttgtcgtatg	cttttgataaa	960	
gatgatggag	gaatacatgg	cgatgcgtca	gagtatgcaa	caggagaaaa	ttttgatact	1020	
gtagtgttcc	acaagaggga	agaacaaaaag	accaatggga	gtgccagtaa	gaagaggcgt	1080	
ctcacgcctg	acactagtaa	tatgggaaca	agcactgatg	tgcaagaatt	ccaacgatg	1140	
ggaacaaata	ctgatatgca	agaattccaa	tcaatgggaa	caaataccaa	ccccatagag	1200	
acttctctag	tgggtgtgaa	taccaaccca	cttcccaacc	ctcccccaag	attggttaatt	1260	
actcctttaa	cgaatgatgt	accagaattg	gacatgatgt	ggcttttatc	gctctccaga	1320	
ggaggtggaa	attctagaat	gagtgcaaat	acaggaacat	ctccccgtgc	taacacccca	1380	
attcctacct	gcttcacagg	aggtgcaaat	gtagtagtgc	ctaattggatt	tgtccctccc	1440	
acgttcccct	tagaattgtg	cgaagatgat	ccaagtattc	ccaattctta	caattacgaa	1500	
gaggataaag	tctttcatcc	attttatgag	tatatggcca	aatatctatc	ccctcttggt	1560	
ccatcatata	acaagggaca	gacttghtaat	gtgtgtccagg	agtgtgtcaa	gggatccttc	1620	
tctcttgcaa	agcgtagagg	aacagtcccc	aaattctgta	gtaacatttc	ccaagctttc	1680	
ttttgtaata	tggatgtatg	tactgccatg	tgcaaatggg	cgaagactgt	aattagacat	1740	
ggacaatatt	gtaatagatg	tatcgtaaag	aggtcattgt	catccatgct	cgcatatcac	1800	
tacattgttt	gcagagacgc	ttcatgtgat	gttcccaagt	gcagggaaag	ggttcgcaac	1860	
gacatgqatg	actga					1875	

```
<210> 201
<211> 620
<212> PRT
<213> SHRIMP
```

<400> 201  
Met Ser Ser Thr Thr Ser Pro Ser Ser Leu Trp Asp Asp Asp Asp Asp  
1 5 10 15





<210>	202
<211>	2820
<212>	DNA
<213>	SHRIMP

atgtctgagc	cttcagttta	tgctttttatt	gacataaaaag	aaatagaaaa	tgggtgggaa	60
aaagaattcg	ggctttttagt	acaaccagga	cagaaattag	ctcctttcag	ggatatttct	120
tatgactcaa	gcaaacttga	ttgtgacgca	ttttcttgca	taccttcaga	tatacttcatt	180
tctgataatg	aaaaaagatg	aggagagtgc	aactttggcg	aacacacctc	tgtctcattt	240
ccgtgcaaga	accctgaggg	aaaaaacattg	gcgcatttca	cggcatgtgg	tccaggggtg	300
taccggaggt	acaagcaaag	agacccccat	actggtttgc	cagtagccag	aggcgctctg	360
atgcaagatc	acgttgacca	tgaactgga	aataaaatgt	gtgaatatct	gaaccagagc	420
ttagtcatgt	gggctgctgt	cccttggata	cgacctggag	accttactga	aggttacaac	480
acaacacacg	tgcttggctt	tgcattcaag	gaggatgacg	aaaggagttc	aaaaagagtt	540
aaatatgaaa	atgtggctct	tcaaaggct	tattgtgatt	tctttaaaca	gtattatgac	600
cgacactctg	gctcatgtta	tcgatctgga	tggatgaaat	ttgtccattt	aatgttttgg	660
cagtatttta	ctaattcttc	ttacaattta	gctaacccaa	aaccttacaa	tttaactgga	720
aatacatggt	ctgatgtagt	ttctgtattg	acagataaacc	ctattgtaga	tgcggggtgct	780
gctccgtcca	gatctgaaat	ggatgaaat	atcaccaaga	agaagtcca	cgtattccct	840
tctgagcaga	catctgctcg	ccagaagaac	gagaataata	tacgtttcca	gtacggagat	900
gggtgagaaa	tagaccctac	ttctgtggat	gctttaatgc	agtttgttaa	tagggaaggt	960
gtagtaggaa	cagagaaaaa	atccgaccgc	ctcatgcgag	tggcagacgc	tgttatggat	1020
gcggctatgc	gccttcaagt	catgggtctg	gacgacagtc	aatctagacg	attattgtta	1080
aaaaatatga	ttaaaatgag	tagaaacaac	ccagaatatg	caagacattt	ttccagttct	1140
ctcaaattaa	ttggagtaac	tttggccata	aaaagactg	ttttctctaa	aggcgcttca	1200
gctaaaagaa	aggaaacggc	cattaataat	ggcgaaacgc	atagaagaag	cagatgggtca	1260
cccgagacgg	tgacagaaga	agatgcactt	ttatttgcaa	gagaaaatat	caactgaagac	1320
ccaaaacacc	ctgcgcctct	tgtggacatt	ttacactcac	cagatatcaa	ttcatctatc	1380
aagagcgggt	cttcttctct	aatatggaac	gatatacttt	caagaatttc	ttcgaccaga	1440
aaactggaag	aaaaggcgag	tgttttcggt	aaaaatcttg	tggtgaaagt	agtgagacag	1500
ttcttggaac	ttttagaagg	ttaactgttt	tcggacggat	accgaatgga	cgataatat	1560
cctctgatga	taggcgtgga	ccaaatattg	agagaagtca	ttaaggcggc	cagtaatatg	1620
tgcgctagat	ttgcctcctc	agctctggag	tctagtttgg	ttactggctt	tatagactcg	1680
gcgagtgcta	tcacttctag	gttagccgta	caactggcag	ccagaacggt	ctccgtattc	1740
ttggaggagt	ctgttataga	atttgtagtt	gccgcaagcc	ttcggctagc	gatacaggga	1800
tttgccgcat	tgccaactct	tgccgcgtca	gccttaaccg	tatttggaat	tgttatatct	1860
gtaatacaag	tacttgggtt	gattttgcat	cttcgcgtag	ggttagggtg	gtacgatcac	1920
atTTTTtagcc	cagaggattt	aaagaagcag	gttctagtgt	ttaggagaga	gtttgcaaag	1980
gcaggaaatg	tagatgtggg	tgtagctcaa	ccagtcaccc	ctgaagaaat	cgctcgctatc	2040
aacgttttcc	ttcaaactga	agaaaacggg	gaagaaaaga	aggaagaagg	agcacagaaa	2100
tcaaagattg	atTTTcttca	aaaatacttc	cattctactc	ctttgatggg	aaagaaaaagt	2160
aagtttgtct	acatacaaga	agcagctcaa	gaatacttgg	gaggaagaac	aatgaacgct	2220
tttgggcagc	gtataataac	agctgctgat	gatagtgaca	ccaccaccac	cacacaagag	2280

```

ggaaggaggg atgacgaaac agtgactaag aaaatgagga gtattattct agaaacaggt 2340
caaactctta aggattactc gtctgctgtt aactataacg cctcccgctc agattacgtg 2400
ggagaggaat gggtaagaaa tactgcccta aaagaagaga caaggagcaa cactactagt 2460
gataacctat tcaagaaaac tgtttctctt gctagtatgg ccggcgcatt tctggttcta 2520
ggaatagggtg tattggtagc gtcccatatt acacttttac gtttcaccaa tattggtcta 2580
gcttttgctg tgcgaggctc cctagcattt attgcactta tgagtatatc atataaaac 2640
atgaatgcta tgggtgtagt gaattcggac gcaatataca ggtctactgc tctagttgga 2700
gatatcaaaa cagaccccag aagagtagga atgggccagc gccacgtagg tgtcggggct 2760
aaatacaaca tgattacaga tttcgtctct ccaatgtagg acgagatcga gagtactaa 2820

```

<210> 203  
 <211> 931  
 <212> PRT  
 <213> SHRIMP

<400> 203

Met	Ser	Glu	Pro	Ser	Val	Tyr	Ala	Phe	Ile	Asp	Ile	Lys	Glu	Ile	Glu	1	5	10	15
Asn	Gly	Trp	Glu	Lys	Glu	Phe	Gly	Leu	Leu	Val	Gln	Pro	Gly	Gln	Lys	20	25	30	
Leu	Ala	Pro	Phe	Arg	Asp	Ile	Ser	Tyr	Asp	Ser	Ser	Lys	Leu	Asp	Cys	35	40	45	
Asp	Ala	Phe	Ser	Cys	Ile	Pro	Ser	Asp	Ile	Leu	His	Ser	Asp	Asn	Glu	50	55	60	
Lys	Arg	Val	Gly	Glu	Cys	Asn	Phe	Ala	Glu	His	Thr	Ser	Val	Ser	Phe	65	70	75	80
Pro	Val	Lys	Asn	Pro	Glu	Gly	Lys	Thr	Leu	Arg	His	Phe	Thr	Ala	Cys	85	90	95	
Gly	Pro	Gly	Cys	Tyr	Arg	Arg	Tyr	Lys	Gln	Arg	Asp	Pro	His	Thr	Gly	100	105	110	
Leu	Pro	Val	Arg	Val	Leu	Met	Gln	Asp	His	Val	Asp	His	Glu	Thr	Gly	115	120	125	
Asn	Lys	Met	Cys	Glu	Tyr	Leu	Asn	Gln	Ser	Leu	Val	Met	Trp	Ala	Ala	130	135	140	
Val	Pro	Trp	Ile	Arg	Pro	Gly	Asp	Leu	Thr	Glu	Gly	Tyr	Asn	Thr	Thr	145	150	155	160
His	Val	Pro	Gly	Phe	Ala	Phe	Lys	Glu	Asp	Asp	Glu	Arg	Asp	Ser	Lys	165	170	175	
Arg	Val	Lys	Tyr	Glu	Asn	Val	Val	Ile	Ser	Lys	Ala	Tyr	Cys	Asp	Phe	180	185	190	
Phe	Lys	Gln	Tyr	Tyr	Asp	Ala	Asp	Ser	Gly	Ser	Cys	Tyr	Arg	Ser	Gly	195	200	205	
Trp	Met	Lys	Phe	Val	His	Leu	Met	Phe	Gly	Gln	Tyr	Phe	Thr	Asn	Leu	210	215	220	
Ser	Tyr	Asn	Leu	Ala	Asn	Pro	Lys	Pro	Tyr	Asn	Leu	Thr	Gly	Asn	Thr	225	230	235	240
Trp	Ser	Asp	Val	Val	Ser	Val	Leu	Thr	Asp	Asn	Pro	Ile	Val	Asp	Ala	245	250	255	
Gly	Ala	Ala	Pro	Ser	Arg	Ser	Glu	Met	Asp	Glu	Ile	Ile	Thr	Lys	Lys	260	265	270	
Lys	Phe	Asn	Val	Phe	Pro	Ser	Glu	Gln	Thr	Ser	Ala	Arg	Gln	Lys	Ala	275	280	285	
Glu	Asn	Ile	Ile	Arg	Ser	Gln	Tyr	Gly	Asp	Gly	Val	Glu	Ile	Asp	Pro	290	295	300	
Ser	Ser	Val	Asp	Ala	Leu	Met	Gln	Phe	Val	Asn	Arg	Glu	Gly	Val	Val	305	310	315	320
Gly	Thr	Glu	Lys	Lys	Ser	Asp	Arg	Leu	Met	Arg	Val	Ala	Asp	Ala	Val	325	330	335	
Met	Asp	Ala	Ala	Met	Arg	Leu	Gln	Val	Met	Gly	Leu	Asp	Asp	Ser	Gln	340	345	350	

```

Ser Arg Arg Leu Leu Leu Lys Asn Met Ile Lys Met Ser Arg Asn Asn
    355          360          365
Pro Glu Tyr Ala Arg His Phe Ser Ser Ser Leu Lys Leu Ile Gly Val
    370          375          380
Thr Leu Ala Ile Lys Arg Ser Val Phe Ser Lys Gly Ala Ser Ala Lys
385          390          395          400
Arg Lys Glu Thr Ala Ile Asn Asn Gly Glu Gln His Arg Arg Ser Arg
    405          410          415
Trp Ser Pro Glu Thr Val Thr Glu Glu Asp Ala Leu Leu Phe Ala Arg
    420          425          430
Glu Asn Ile Thr Glu Asp Pro Lys His Pro Ala Pro Phe Val Asp Ile
    435          440          445
Leu His Ser Pro Asp Ile Asn Ser Ser Ile Lys Ser Gly Ser Ser Ser
    450          455          460
Ser Ile Trp Asn Asp Ile Leu Ser Arg Ile Ser Ser Thr Arg Lys Leu
465          470          475          480
Glu Glu Lys Ala Ser Val Phe Val Lys Asn Leu Val Val Lys Val Val
    485          490          495
Arg Gln Phe Leu Asp Ile Gly Lys Leu Phe Ser Asp Gly Tyr Glu Trp
    500          505          510
Asp Asp Asn Ile Pro Leu Met Ile Gly Val Asp Gln Ile Leu Arg Glu
    515          520          525
Val Ile Lys Ala Asn Met Cys Ala Arg Phe Ala Ser Ser Ala Leu Glu
    530          535          540
Ser Ser Leu Val Thr Gly Phe Ile Asp Ser Ala Ser Ala Ile Thr Ser
545          550          555          560
Arg Leu Ala Val Gln Leu Ala Ala Arg Thr Phe Ser Val Phe Leu Glu
    565          570          575
Glu Ser Val Ile Glu Phe Val Val Ala Ala Ser Leu Arg Leu Ala Ile
    580          585          590
Gln Ala Phe Ala Asp Leu Ala Thr Leu Ala Ala Ser Ala Leu Thr Val
    595          600          605
Ile Gly Ile Val Ile Phe Val Ile Gln Val Leu Gly Leu Ile Leu Asp
    610          615          620
Leu Ala Leu Gly Leu Gly Trp Tyr Asp His Ile Phe Ser Pro Glu Asp
625          630          635          640
Leu Lys Lys Gln Val Leu Val Phe Arg Arg Glu Phe Ala Lys Ala Gly
    645          650          655
Asn Val Asp Val Gly Val Ala Gln Pro Val Thr Pro Glu Glu Ile Val
    660          665          670
Ala Ile Asn Val Phe Leu Gln Thr Glu Glu Asn Gly Glu Glu Lys Lys
    675          680          685
Glu Glu Gly Ala Arg Lys Ser Lys Ile Asp Phe Leu Gln Lys Tyr Phe
    690          695          700
His Ser Thr Pro Leu Met Gly Lys Lys Ser Lys Phe Val Tyr Ile Gln
705          710          715          720
Glu Ala Ala Gln Glu Tyr Leu Gly Gly Arg Thr Met Asn Ala Phe Gly
    725          730          735
Gln Arg Ile Ile Thr Ala Ala Asp Asp Ser Asp Thr Thr Thr Thr Thr
    740          745          750
Gln Glu Gly Arg Arg Asp Asp Glu Thr Val Thr Lys Lys Met Arg Ser
    755          760          765
Ile Ile Thr Gly Gln Thr Leu Lys Asp Tyr Ser Ser Ala Val Asn Tyr
    770          775          780
Asn Ala Ser Arg Leu Asp Tyr Val Gly Glu Glu Trp Val Arg Asn Thr
785          790          795          800
Ala Leu Lys Glu Glu Thr Arg Ser Asn Thr Thr Ser Asp Asn Leu Phe
    805          810          815
Lys Lys Thr Val Ser Leu Ala Ser Met Ala Gly Ala Phe Leu Val Leu
    820          825          830
Gly Ile Gly Val Leu Val Ala Ser His Ile Thr Leu Leu Arg Phe Thr

```

835	840	845
Asn Ile Gly Leu Ala Phe	Ala Phe Ala Gly Leu	Leu Ala Phe Ile Ala
850	855	860
Leu Met Ser Ile Ser Tyr	Ile Asn Met Asn Ala	Met Gly Val Val Asn
865	870	875
Ser Asp Ala Ile Tyr Arg	Ser Thr Ala Leu Val	Gly Asp Ile Lys Thr
885	890	895
Asp Pro Arg Arg Val Gly	Met Val Gln Arg His	Val Gly Val Gly Ala
900	905	910
Lys Tyr Asn Met Ile Thr	Asp Phe Val Ser Pro	Met Leu Asp Glu Ile
915	920	925
Glu Ser Asp		
930		

<210> 204  
 <211> 3135  
 <212> DNA  
 <213> SHRIMP

<400> 204  
 atggattcgt tgactaatac agtaacactc ctcgtgaatg accgtcttgg aaatcatcga 60  
 acaaacaaac caatcaccga acaagatgtg gaaaatactc ttaaccttaa cagtctagaa 120  
 agggcaagtc ttttgaagtt atattctgtt ttcatacaag aaatgcagtc ctattctgga 180  
 tgtataccca agaacaagta cacgaatgtg caagaaatat tcgaagatgg actaattact 240  
 ttcgaatgga gagatggaac aaaagtacac agatcagttt caccaagttc ccctatacct 300  
 ctttctacaa aaaaatcgcc tcggtcctca ccttccccctc ctccatcgat gccttctatt 360  
 aaagaagaag aatttgaaga agaatttgaa gatgatgaag aaatatacga aacagatgaa 420  
 aatgtggaag atttcataaa tggatgatgga gaagattcag aagaggaaga agaagaagat 480  
 ataattgttg atgacgaaga agaagagaat gaagagggag aaaacaagta cgttctcgca 540  
 ttctctaate atctgaggcg ccagactgcc gccgctgccg ccgcccgcgc agctgctgct 600  
 gctgctgaca ttgagaagaa ggacaaaaaac cagcagttta gtgcatga ctacacactg 660  
 tccgccctcc agcaacagca acaaaaactg ctccagcaac agcaacagca acaacaccag 720  
 cagcgctcct catctgagaa ggtcacctcc acaccaaca aattcaacaa gttttactg 780  
 ccgagtaatg gcttctcaga acagaccgag ctctttgttt gtttcgatgt ggataaaatt 840  
 gcgcaatata atggactcgt ggagctagac atcttaccga ttgttgctga atacattatc 900  
 aacggccttg gtctgaaatg cagtatggaa actccccag tgaaaccgtg caggagaaag 960  
 gaagtgaag atgtgtggtg tcagcctaaa actagctttg aaaatgatgc tgtggaagat 1020  
 aaacatctcg cattcgcaga atcgccataa cttcaaaggc ctagagattt ccctatccct 1080  
 aaaaaaatca ccgcctattt ttgttttagac gattctgtag acattaaaaa cccctggggg 1140  
 tcgtgtccgc ttttgaaaag tggatcaaac ttccgagtgt ccgaatattc gcgtcatttt 1200  
 aatgaattct caggagttaa aaatgacgat gacacgtctt caaacacttg ctttatatac 1260  
 tctcaaaaaa accccaacat tgaaattgta tcaaaattaa atattgaatt tgaggtaatg 1320  
 atggagggaa ttataaccca tagaaaagat ttgttcgaga cgggcatttt gagcgattct 1380  
 tcattagcta cggctatggc attttgccac ccaaagcta gagttcgaaa tgttgcatgt 1440  
 ttttattttt ctgtatattt accattctctt aaaaataactc gcaaagaaac tataaaattg 1500  
 tcagagacgg ataaggtaca tattggttca gacgcgatct tttctcccc aagtgataat 1560  
 cctaataataa gtgctcacca gaacaataac acaataataa acaataatac cagtgtgaat 1620  
 atcgaggaca gacctatccg aaataataat ataagcagaa aaatgaccat cacaaactac 1680  
 caatgtatgg catgcaagga aagatgcaca aacaattgca ctaacggtaa ctatcccgat 1740  
 cgtggttaacc agcacttgct acatagtgtg aaaggggaag atttctttta gatttttaaa 1800  
 aatagtaaa tagattcatt gaaaaaattg agcagagtac tgattccgc tctccctct 1860  
 ggaaattata catctaagtt ttgtgataga agctctatgt gccatagctt ctttttaga 1920  
 gggatcgaac cagtgtctac ttctttctca tcggacagtt ttgaaaagac taaacttggt 1980  
 ttgtatggta aagtcgttga cgttatcaat agttattctg ccataaaaaac ttcccataat 2040  
 aatagaatca gggcttttt taactctgag gaaaaagata ataagactat cccctctaga 2100  
 gccgaaagtg caaaaaatgc attcaaggat atacttggtc acgaatgtaa taaagaacga 2160  
 gctgtttcat attttgagca aaacaaatta tcctctaaag atgggcatct atctaacaag 2220  
 tgggtggatcg aacttaatga cttgaacatt atgtttgaga aacacgtgga agatttttac 2280  
 aagaaatgtt ctaaagtaaa tgatgcagaa tcttttaaagg atatttttaa tgattttgaa 2340  
 aaaacttggt ataaatacaa aactgccaag agggcaatta ttggagcaca agacccttct 2400

```

acttctactc cctctaaaaa ggagaatggt atcactagga ttattagtagc attatccgaa 2460
tttcattcaa aagatgaagc tacagtaagt gcccttctcg acaaaacaat gctcttggga 2520
tcgaggacaa taatgtctgg tgtagatgt gttatacgta acaatagtgt gttttcgggc 2580
tttgaataa agaacactaa taataattgg gaacttgaga ttagacacta tgcattctct 2640
atgggaggtg ctgcagtgc aaagatttcc gatgaagatt tggaacaatt cacgcctgta 2700
agaggtgctg tctctgtcac tacagcacct aatgataagc tacctgtagg ggcacatcag 2760
acatggaagg atgaacaaac actaaaaaca aacactaaac gtaatagtct atatgactct 2820
tacaattcaa aaaggaataa tagggataat aataaaataa aaaatcggtc attaaaacta 2880
tcagatttta attggagaac acccaatata tctattcaag aatttaaatgc aaataaagat 2940
gatgttaaca agaagaggta cgcagaagtc gtggcgtagc ctgctccaaa gtcaccttca 3000
ccaacgagca gcagcagcag caacagcaac agcagcagcc ctcctctttc accgctctca 3060
ccaacagtga agaataagtaa taataaacca ttgtatatcc ctccccataa aagaatgacc 3120
actactgctg tttga                                     3135

```

<210> 205  
 <211> 1036  
 <212> PRT  
 <213> SHRIMP

<400> 205

Met	Asp	Ser	Leu	Thr	Asn	Thr	Val	Thr	Leu	Leu	Val	Asn	Asp	Arg	Leu
1				5					10					15	
Gly	Asn	His	Arg	Thr	Asn	Lys	Pro	Ile	Thr	Glu	Gln	Asp	Val	Glu	Asn
			20					25					30		
Thr	Leu	Asn	Leu	Asn	Ser	Leu	Glu	Arg	Ala	Ser	Leu	Leu	Lys	Lys	Val
			35				40					45			
Phe	Ile	Lys	Glu	Met	Gln	Ser	Tyr	Ser	Gly	Cys	Ile	Pro	Lys	Asn	Lys
	50				55					60					
Tyr	Thr	Asn	Val	Gln	Glu	Ile	Phe	Glu	Asp	Gly	Leu	Ile	Thr	Phe	Glu
65				70					75					80	
Trp	Arg	Asp	Gly	Thr	Lys	Val	His	Arg	Ser	Val	Ser	Pro	Ser	Ser	Pro
				85				90						95	
Ile	Pro	Leu	Ser	Thr	Lys	Lys	Ser	Pro	Arg	Ser	Ser	Pro	Ser	Pro	Pro
			100					105					110		
Pro	Ser	Met	Pro	Ser	Ile	Lys	Glu	Glu	Phe	Glu	Glu	Glu	Phe	Glu	
		115				120						125			
Asp	Asp	Glu	Glu	Ile	Tyr	Glu	Thr	Asp	Glu	Asn	Val	Glu	Asp	Phe	Ile
	130					135					140				
Asn	Gly	Asp	Gly	Glu	Asp	Ser	Glu	Glu	Glu	Glu	Glu	Glu	Asp	Ile	Ile
145					150					155					160
Val	Asp	Asp	Glu	Glu	Glu	Asn	Glu	Glu	Gly	Glu	Asn	Lys	Tyr	Val	
			165					170					175		
Leu	Ala	Phe	Ser	Asn	His	Leu	Arg	Arg	Gln	Thr	Ala	Ala	Ala	Ala	Ala
		180						185					190		
Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Asp	Ile	Glu	Lys	Lys	Asp	Lys	Asn
		195					200					205			
His	Ala	Val	Ser	Ala	His	Asp	Tyr	Thr	Leu	Ser	Ala	Leu	Gln	Gln	Gln
	210					215					220				
Gln	Gln	Lys	Leu	Leu	Gln	Gln	Gln	Gln	Gln	Gln	Gln	His	Gln	Gln	Arg
225					230					235					240
Ser	Ser	Ser	Glu	Lys	Val	Thr	Ser	Thr	Pro	Asn	Lys	Phe	Asn	Lys	Phe
				245					250					255	
Leu	Leu	Pro	Ser	Asn	Gly	Phe	Ser	Glu	Gln	Thr	Glu	Leu	Phe	Val	Cys
		260						265					270		
Phe	Asp	Val	Asp	Lys	Ile	Ala	Gln	Tyr	Asn	Gly	Leu	Val	Glu	Leu	Asp
	275					280					285				
Ile	Leu	Pro	Ile	Val	Ala	Glu	Tyr	Ile	Ile	Asn	Gly	Leu	Gly	Leu	Lys
	290					295					300				
Cys	Ser	Met	Pro	Pro	Val	Lys	Pro	Cys	Arg	Arg	Lys	Glu	Val	Lys	Asp
305					310					315					320
Val	Trp	Cys	Gln	Pro	Lys	Thr	Ser	Phe	Glu	Asn	Asp	Ala	Val	Glu	Asp

Lys	His	Leu	Ala	Phe	Ala	Glu	Ser	Pro	Ile	Leu	Gln	Arg	Pro	Arg	Asp	
			340					345					350			
Phe	Pro	Ile	Pro	Lys	Lys	Ile	Thr	Ala	Tyr	Phe	Cys	Leu	Asp	Asp	Ser	
			355				360					365				
Val	Asp	Ile	Lys	Asn	Pro	Trp	Gly	Ser	Cys	Pro	Leu	Leu	Lys	Ser	Gly	
			370			375					380					
Ser	Asn	Phe	Arg	Val	Ser	Glu	Tyr	Ser	Arg	His	Phe	Asn	Glu	Phe	Ser	
385					390					395					400	
Gly	Val	Lys	Asn	Asp	Asp	Asp	Thr	Ser	Ser	Asn	Thr	Cys	Phe	Ile	Tyr	
				405					410					415		
Ser	Gln	Lys	Asn	Pro	Asn	Ile	Glu	Ile	Val	Ser	Lys	Leu	Asn	Ile	Glu	
			420					425					430			
Phe	Glu	Val	Met	Met	Glu	Gly	Ile	Ile	Thr	Lys	Asp	Leu	Phe	Glu	Thr	
			435				440					445				
Gly	Ile	Leu	Ser	Asp	Ser	Ser	Leu	Ala	Thr	Ala	Met	Ala	Phe	Cys	His	
			450			455					460					
Pro	Lys	Ala	Arg	Val	Arg	Asn	Val	Phe	Tyr	Phe	Ser	Val	Tyr	Leu	Pro	
465					470					475					480	
Phe	Ser	Lys	Ile	Thr	Arg	Lys	Glu	Thr	Ile	Lys	Cys	Ser	Glu	Thr	Asp	
				485					490					495		
Lys	Val	His	Ile	Gly	Ser	Asp	Ala	Ile	Phe	Ser	Pro	Pro	Ser	Asp	Asn	
			500					505					510			
Pro	Asn	Ile	Ser	Ala	His	Gln	Asn	Asn	Asn	Asn	Asn	Asn	Asn	Asn	Asn	
			515				520						525			
Thr	Ser	Val	Asn	Ile	Glu	Asp	Arg	Pro	Ile	Arg	Asn	Asn	Asn	Ile	Ser	
			530			535					540					
Arg	Lys	Met	Thr	Ile	Thr	Asn	Tyr	Gln	Cys	Met	Ala	Cys	Lys	Glu	Arg	
545					550					555					560	
Cys	Thr	Asn	Asn	Cys	Thr	Asn	Gly	Asn	Tyr	Pro	Asp	Arg	Gly	Asn	Gln	
				565					570					575		
His	Leu	Ser	His	Ser	Val	Lys	Gly	Glu	Asp	Phe	Phe	Lys	Ile	Leu	Asn	
			580					585					590			
Asn	Ser	Lys	Val	Asp	Ser	Leu	Lys	Lys	Leu	Ser	Arg	Val	Leu	Ile	Pro	
			595				600					605				
Ala	Pro	Pro	Ser	Gly	Asn	Tyr	Thr	Ser	Lys	Phe	Cys	Asp	Arg	Ser	Ser	
			610			615					620					
Met	Cys	His	Ser	Phe	Phe	Cys	Arg	Gly	Ile	Glu	Pro	Val	Ser	Thr	Ser	
625					630					635					640	
Phe	Ser	Ser	Asp	Ser	Phe	Glu	Lys	Thr	Lys	Leu	Val	Leu	Tyr	Gly	Lys	
				645					650					655		
Val	Val	Asp	Val	Ile	Asn	Ser	Tyr	Ser	Ala	Ile	Lys	Thr	Ser	His	Asn	
			660					665								

WO 01/38351

291

PCT/US00/28888

Asp	Glu	Ala	Thr	Val	Ser	Ala	Leu	Leu	Asp	Lys	Thr	Met	Leu	Leu	Gly
			820					825					830		
Ser	Arg	Thr	Ile	Met	Ser	Gly	Val	Arg	Cys	Val	Ile	Arg	Asn	Asn	Ser
		835					840						845		
Val	Phe	Ser	Gly	Phe	Glu	Asn	Lys	Asn	Thr	Asn	Asn	Asn	Trp	Glu	Leu
	850					855					860				
Glu	Ile	Arg	His	Tyr	Val	Ile	Ser	Met	Gly	Gly	Ala	Ala	Val	Thr	Lys
865					870					875					880
Ile	Ser	Asp	Glu	Asp	Leu	Glu	Gln	Phe	Thr	Pro	Val	Arg	Gly	Ala	Val
				885					890					895	
Ser	Val	Thr	Thr	Ala	Pro	Asn	Asp	Lys	Leu	Pro	Val	Gly	Ala	His	Gln
			900					905					910		
Thr	Trp	Lys	Asp	Glu	Gln	Thr	Leu	Lys	Thr	Asn	Thr	Lys	Arg	Asn	Ser
		915					920					925			
Leu	Tyr	Asp	Ser	Tyr	Asn	Ser	Lys	Arg	Asn	Asn	Arg	Asp	Asn	Asn	Lys
	930					935					940				
Ile	Lys	Asn	Arg	Ser	Leu	Lys	Leu	Ser	Asp	Phe	Asn	Trp	Arg	Thr	Pro
945					950					955					960
Asn	Ile	Ser	Ile	Gln	Glu	Phe	Asn	Ala	Asn	Lys	Asp	Asp	Val	Asn	Lys
				965					970					975	
Lys	Arg	Tyr	Ala	Glu	Val	Val	Ala	Ser	Ala	Ala	Pro	Lys	Ser	Pro	Ser
			980					985					990		
Pro	Thr	Ser	Ser	Ser	Ser	Ser	Asn	Ser	Asn	Ser	Ser	Ser	Pro	Pro	Leu
		995					1000					1005			
Ser	Pro	Leu	Ser	Pro	Thr	Val	Lys	Asn	Ser	Asn	Asn	Lys	Pro	Leu	Tyr
	1010					1015					1020				
Ile	Pro	Pro	His	Lys	Arg	Met	Thr	Thr	Thr	Ala	Val				
1025					1030					1035					

<210> 206  
 <211> 1082  
 <212> DNA  
 <213> SHRIMP

<400> 206  
 atgaagatat gccagatttc ttcccctacc ctcaccctga gcatccatt ggagggcgtc 60  
 taccacgtga agcagctcct ccacttaaag gtgcacttgg acgtaaaagg cgtgaagcag 120  
 ctccctccact taaagggtgcg cttggacgta agaggcgca agcagaatcc ttggaggaag 180  
 aacttgtgtc tgctgaagaa gaacgtgaaa agcgcgagc agtccccca cttaaagggtg 240  
 cacttggacg tgaaaagcgc gaagcagctc cccacttaa aggtgcactt ggacgtaaga 300  
 ggcgcgaagc agtccccca cttaaagggtg cgcttggacg tgaaaagcgc gaagcagctc 360  
 cccacttaa aggtgcactt ggacgtaaga ggcgcgaagc agtccccca cttaaagggtg 420  
 cgcttggacg taagaggcgc gaagcagaat cttggagga agaacttgtg tctgtgaag 480  
 aagaacgtga aaagcgcgaa gcagctcccc cacttaaagg tgcacttggc cgtaaaaggc 540  
 gtgaagcagc tcctccactt aaagggtgcg cttggacgta gaggcgcgaa gcagctcccc 600  
 cacttaaagg tgcacttggc cgtaagaggc gcgaagcaga atccttggag gaagaacttg 660  
 tgtctgttga agaagaacgt gaaaagcgcg aagcagctcc cccacttaaa ggtgctcttg 720  
 gacgtaagag gcgcgaagca gctccccac ttaaagggtg ctttggacgt aagaggcgcg 780  
 aagcagctcc cccacttaaa ggtgctcttg gacgtaagag gcgcgaagca gaatccttgg 840  
 aggaagaact tgtgtctgct gaagaagaac gtgaaaagcg cgaagcagct cccccactta 900  
 aagggtgctct tggacgtaag aggcgcgaag cagctcccc acttaaagggt gcacttggac 960  
 gtaagaggcg cgaagcagca gcagcagcta tgcctcccc tgaagacgat ctgcacttct 1020  
 ttacgcacc tggtgtcttg cctctacatg gagtatggaa agcaccagaa cctacaggtt 1080  
 aa 1082

<210> 207  
 <211> 361  
 <212> PRT  
 <213> SHRIMP

<220>  
 <221> VARIANT  
 <222> (1)...(361)  
 <223> Xaa = Any Amino Acid

<400> 207

Met	Lys	Ile	Cys	Gln	Ile	Ser	Ser	Pro	Thr	Leu	Thr	Leu	Ser	Ile	Pro
1			5						10					15	
Leu	Glu	Gly	Val	Tyr	His	Val	Lys	Gln	Leu	Leu	His	Leu	Lys	Val	His
			20					25					30		
Leu	Asp	Val	Lys	Gly	Val	Lys	Gln	Leu	Leu	His	Leu	Lys	Val	Arg	Leu
	35						40					45			
Asp	Val	Arg	Gly	Ala	Lys	Gln	Asn	Pro	Trp	Arg	Lys	Asn	Leu	Cys	Leu
	50					55					60				
Leu	Lys	Lys	Asn	Val	Lys	Ser	Ala	Lys	Gln	Leu	Pro	His	Leu	Lys	Val
65					70					75					80
His	Leu	Asp	Val	Lys	Ser	Ala	Lys	Gln	Leu	Pro	His	Leu	Lys	Val	His
			85						90					95	
Leu	Asp	Val	Arg	Gly	Ala	Lys	Gln	Leu	Pro	His	Leu	Lys	Val	Arg	Leu
			100					105					110		
Asp	Val	Lys	Ser	Ala	Lys	Gln	Leu	Pro	His	Leu	Lys	Val	His	Leu	Asp
	115						120					125			
Val	Arg	Gly	Ala	Lys	Gln	Leu	Pro	His	Leu	Lys	Val	Arg	Leu	Asp	Val
	130					135					140				
Arg	Gly	Ala	Lys	Gln	Asn	Pro	Trp	Arg	Lys	Asn	Leu	Cys	Leu	Leu	Lys
145					150					155					160
Lys	Asn	Val	Lys	Ser	Ala	Lys	Gln	Leu	Pro	His	Leu	Lys	Val	His	Leu
			165						170					175	
Asp	Val	Lys	Gly	Val	Lys	Gln	Leu	Leu	His	Leu	Lys	Val	Arg	Leu	Asp
			180					185					190		
Val	Arg	Gly	Ala	Lys	Gln	Leu	Pro	His	Leu	Lys	Val	His	Leu	Asp	Val
	195						200					205			
Arg	Gly	Ala	Lys	Gln	Asn	Pro	Trp	Arg	Lys	Asn	Leu	Cys	Leu	Leu	Lys
	210					215					220				
Lys	Asn	Val	Lys	Ser	Ala	Lys	Gln	Leu	Pro	His	Leu	Lys	Val	Leu	Leu
225					230					235					240
Asp	Val	Arg	Gly	Ala	Lys	Gln	Leu	Pro	His	Leu	Lys	Val	Leu	Leu	Asp
			245						250					255	
Val	Arg	Gly	Ala	Lys	Gln	Leu	Pro	His	Leu	Lys	Val	Leu	Leu	Asp	Val
			260					265					270		
Arg	Gly	Ala	Lys	Gln	Asn	Pro	Trp	Arg	Lys	Asn	Leu	Cys	Leu	Leu	Lys
	275						280				285				
Lys	Asn	Val	Lys	Ser	Ala	Lys	Gln	Leu	Pro	His	Leu	Lys	Val	Leu	Leu
	290					295					300				
Asp	Val	Arg	Gly	Ala	Lys	Gln	Leu	Pro	His	Leu	Lys	Val	His	Leu	Asp
	305				310					315					320
Val	Arg	Gly	Ala	Lys	Gln	Gln	Gln	Gln	Leu	Cys	Leu	Pro	Leu	Lys	Thr
			325					330						335	
Ile	Ser	Thr	Ser	Phe	Thr	His	Leu	Leu	Leu	Cys	Leu	Tyr	Met	Glu	Tyr
			340					345					350		
Gly	Lys	His	Gln	Asn	Leu	Gln	Val	Xaa							
		355					360								

<210> 208  
 <211> 816  
 <212> DNA  
 <213> SHRIMP

<400> 208  
 atggacgtcg agttcgggttt cttccacggt ctgctctcca aggcccttct cccagatgaa 60



```

aaacatcaac cccgttataag gcgtctttgt gcggatgatt ctagaaataa gggagaggat 120
ggctgctgct cgttctgtgg aagaagagga acaggagaga gcaatactgc gtgccttgaa 180
caactaatag acgtttgttc ctttatagga actgtctcat ctattggtac aattatcaat 240
tctaattctt caaccagttg ttctagacta caaaaaacgt cagacagtta tgcggcatta 300
tcccattcta gttttctgga tgtggtatat ccaagtttga agaaaacaac tgaagacgta 360
ttgcctcatt ctttacgtgc catttggaat aaacaacttc caaagttgta tgaaaaaact 420
cttcaaccca tagaagaaga ggatattggg tataaggatt atgttggttc aattgaagac 480
gacgacaatg ttgatgatgg tgaccaacaa gaacaaatga ttattgatga agaattctat 540
aaaactattg gagaaaaatc aaccattgaa ctgataggca tgtataacaa taacaagttt 600
ggtaatgaat ttataaggat tcctttaaga gaaactgcgt tgcacgcaca atctctgagg 660
tacgacactg aagctaaatt tgtaaccac aaggactcta tacctctatt ttatgaaaaac 720
agcacgtgca catgtaagga acgtcttatt gatttttctg agagacaact acaacaacta 780
aaacaagatg gaatggataa accaacggac aagtag 816

```

<210> 209  
 <211> 269  
 <212> PRT  
 <213> SHRIMP

<400> 209

```

Met Asp Val Glu Phe Gly Phe Phe His Gly Leu Leu Ser Lys Ala Leu
 1          5          10          15
Leu Pro Asp Glu Lys His Gln Pro Val Ile Arg Arg Leu Cys Ala Asp
          20          25          30
Asp Ser Arg Asn Lys Gly Glu Asp Gly Cys Cys Ser Phe Cys Gly Arg
          35          40          45
Arg Gly Thr Gly Glu Ser Asn Thr Ala Cys Leu Glu Gln Leu Ile Asp
          50          55          60
Val Cys Ser Phe Ile Gly Thr Val Ser Ser Ile Gly Thr Ile Ile Asn
65          70          75          80
Ser Asn Leu Ser Thr Ser Cys Ser Arg Leu Gln Lys Thr Ser Asp Ser
          85          90          95
Tyr Ala Ala Leu Ser His Ser Ser Phe Leu Asp Val Val Tyr Pro Ser
          100          105          110
Leu Lys Lys Thr Thr Glu Asp Val Leu Pro His Ser Leu Arg Ala Ile
          115          120          125
Trp Asn Lys Gln Leu Pro Lys Leu Tyr Glu Lys Thr Leu Gln Pro Ile
          130          135          140
Glu Glu Glu Asp Ile Gly Tyr Lys Asp Tyr Val Val Ser Ile Glu Asp
145          150          155          160
Asp Asp Asn Val Asp Asp Gly Asp Gln Gln Glu Gln Met Ile Ile Asp
          165          170          175
Glu Glu Ser Tyr Lys Thr Ile Gly Glu Lys Ser Thr Ile Glu Leu Ile
          180          185          190
Gly Met Tyr Asn Asn Asn Lys Phe Gly Asn Glu Phe Ile Arg Ile Pro
          195          200          205
Leu Arg Glu Thr Ala Leu His Ala Gln Ser Leu Arg Tyr Asp Thr Glu
210          215          220
Ala Lys Phe Val Asn His Lys Asp Ser Ile Pro Leu Phe Tyr Glu Asn
225          230          235          240
Ser Thr Cys Thr Cys Lys Glu Arg Leu Ile Asp Phe Ser Gln Leu Gln
          245          250          255
Gln Leu Lys Gln Asp Gly Met Asp Lys Pro Thr Asp Lys
          260          265

```

<210> 210  
 <211> 3813  
 <212> DNA  
 <213> SHRIMP

<400> 210

atggtgctaa	ctcttttctg	tacgactagg	cgcgtagcgt	ctagcaaggg	gaattttctct	60
aaggaagatg	cggtgttggg	gaaccagttc	cccattttta	agaaatcaaa	caacttgtca	120
attgcaagac	ctccctcaat	agaatctttt	tctgcatacag	tggaaaaaat	attcagggaa	180
tggaaacgaaa	gtggggggaga	aaaaattttc	gacatatctc	agaatgaaga	agaatggatg	240
gatatcatat	ccttagtggg	aagtgtatat	gaacctgtat	tttctaaatc	acttaaactc	300
gataaatttg	cagataaaac	atgtctaacc	gcggctgcct	ttgcagcact	agcttctgcc	360
gtggatgaaa	aattgacaat	cttatcaggt	agtatggga	gtgtgcttca	acgtacaaca	420
aaggttatga	aaaaggaccc	caaaaaata	gcagaatctc	ttttaaataa	tgaaaaatgg	480
acatctattt	tattggacag	gttaaaaaca	gccaagaaac	ttctaagcag	acgaggtgca	540
ctgaaaagcg	ccgaaagagt	agaagtactt	catcggttga	ataaactcaa	ggaggctcct	600
cttccccacc	atccgagcct	atttgataat	tttagtggag	gaaaaacatc	agcagtatct	660
gctggaacag	tcatacgcatc	agatatgcat	ttcaaatgg	ttgaacatat	ttttaaggtc	720
tccttttagaa	aatgggggtcc	ctgtggagat	aaaactgaaa	gcgggggaaga	agaagatgag	780
gaagaagaag	aagaagaaaa	gaaacatttc	atatcaagat	tcgtgcttca	atttatgaac	840
ggacacaacg	ggcaacatta	tcataggccc	gaaagtgcct	ctgtttactt	ttgtgattat	900
tatgactatt	tggcctacag	gaatctccct	aatgagtaca	aattatcgtc	aatgcatcct	960
ggcacattca	atatggagga	tttacctttc	cgccctttcg	cagtaccttc	aacttataag	1020
acagaattag	agtacaaaag	gtttgtgcaa	tcaacaaatc	ttccccagct	aagtttcgac	1080
tatggggagt	ttttatgtta	ctgtatcttc	ggagcagatt	ggtacaaaaca	cctgggggat	1140
gtggtagatt	ctctagaaaa	tagttccatg	atatcatttg	attctcagac	attgagtggg	1200
gtgtataaga	acactgctaa	ttacaaaagg	ttggggaaga	aaagaaacgg	aatagccgat	1260
ttggccgtta	ggagtatggc	agaattttatc	cgactgaag	cgcataaggc	attgacagca	1320
gaagagatgg	aagaagaaga	agaggaagaa	gaagcggaa	aagaagcgat	ggaccaggag	1380
cctgcagaag	tagacttcc	ttcagtgcc	catttacgcc	gtaaaattcg	tcaagctgtt	1440
tctgtgttaa	ataactttgt	ggagaacgat	ctttctatat	tggtttctaa	cttcaagaat	1500
gtgttaaccg	atgatactgt	atcaggaaca	gatacggaca	actttggttc	tagtggagaa	1560
tttgaagcat	tatcttccca	tttattcctt	tcaagaatat	tggatgaagt	gcacattctt	1620
aggaatactg	atatacaaa	aaccctat	tcaacgcacg	tgtctctgtc	ggataaatct	1680
ccccctagcc	gtgtccgtgg	aagcaatgtc	aactttaata	ataacgctgg	gaacatttct	1740
tccttgcaaa	cgtatggcgg	tatagaagag	ttgcctgaaa	atgtactagt	cggtttgtcc	1800
ggagattttg	aagataccga	catgtattcc	ggagaggatg	ttgttgctgt	atgggatggg	1860
tgtgatggag	gaaaagtgtc	aagtgtcacc	ttcaattgtg	gtgataattt	tatccagctc	1920
catgaaaaaa	cagcagaaac	ttttaaggat	gatacggatt	tagttgaacg	aataagagat	1980
gtgcttcaga	ctgcaagtaa	gaccggaaac	cttaacaaaa	aagcatattc	aaggaagaac	2040
atctacgctg	ttttgcgtga	aaatggcatt	gagcgccccg	gggacgattt	tacagaaaa	2100
gggattgtct	tcaaggataa	aacaaatcaa	ccccctcccc	ctgcaagaag	tgccaagata	2160
acggttgaag	gagtcaaggg	atttttcagc	ggttttcgtg	acattttgga	gacgagggcg	2220
ctcaccacat	atagtgcaga	aaccttcaga	gatttaggcc	aaggcatagt	aaaagagacc	2280
gaaggactga	cagctgcaac	agtggcagaa	acatccttct	ctgaagggtt	agctgaaagt	2340
tttaaggtctg	atgcgaatct	aggtctagaa	ttttcagagg	acgccaaaa	ggttgtatct	2400
aaaaatgaca	cctctcgctt	tttattggaa	gaaactaggg	cattaagagc	aaacaatact	2460
tctttttcgt	cgtttgcaag	ggacatgggc	gtccaagtta	gtgccgattt	agatgctgaa	2520
tttgctgcag	agatgagaga	aacatacccc	gatgcagccc	ttgaacaaaa	cttgaaagat	2580
ctogacaaat	tcgaagagac	tataccagaa	agtcaagtga	agaaactaaa	gaaaatagac	2640
agttatttga	cagagaatcc	agaaaagggt	ggcaaaagaa	ttaacgcac	tgaactgtca	2700
aaggctacag	attcagtatt	ggggaagaaa	caggcaatg	cagttacagt	gttgatgaac	2760
aactttggaa	aggttacaat	tgtagtaggg	gcttctgtgg	tggccgggtt	tttaggtcca	2820
gccgctgtcg	ccctgggtga	tgcgtccaga	ggggcacatc	tcaacgtcgt	ggaccacacc	2880
agccctaaag	gtgtcatcag	ttataaaatt	gtggactttt	cttggtgcaga	tagaaacacc	2940
ggatgggcta	agccaacca	gcacccgttc	agggaaagaa	tagaccatgt	tatcgacta	3000
tatgcatcat	tcttaactga	aaatggagca	tatgtattcc	ctgaagacgg	aggacccaaa	3060
tcgaaatata	agccctacgc	accaatctgt	ggaacaaaag	atgctgctca	aggagaatgt	3120
ggatcttggg	caacattoga	cgacccgcac	tctgtattgc	cttgggtggc	aagcatgaaa	3180
gatttgccta	aaggacaatc	cctctcctgc	gataaaggga	tgtccacttt	aaaggcagtt	3240
tcttccgttc	ttttgtccat	aggaaaggat	gttgacagag	ctatttttga	ggttgcagag	3300
gacgccgtgg	tggggttggc	gagcaaggca	atttcagctg	taataaataa	ccccctgttc	3360
atatttggag	tgccctctgg	atttgggtata	ctgtctacac	gcctcaatcc	atcccaactg	3420
aaaactggcc	tcattgtatt	ttcaatacta	ctagtgttca	tactgatagt	tcgcttcttt	3480
gcagggtcgg	gcccgcatac	cctgaattgg	tttgggtgaa	agaattcagc	taaaaggaaa	3540
cagactgaac	aattcgaaga	cggggggagga	aatcgttcaa	aaatagatt	ggcagaaaa	3600

gacaacgccca atagtaaact tcaatcgagg aggaatgaaa ctggggcccat gagattagag 3660  
 gagcttcctg ggcatagaaga ttgctgcccga gttttcttcc ctgccacaac aaattattcc 3720  
 aaatctgccca agattctggg ctacaaatct aaacccttca acgactttta tacaaaaata 3780  
 ataaacacag acatcataaa aatggatag taa 3813

<210> 211  
 <211> 1264  
 <212> PRT  
 <213> SHRIMP

<400> 211

Met	Val	Leu	Thr	Leu	Ser	Cys	Thr	Thr	Arg	Arg	Val	Ala	Ser	Ser	Lys
1				5					10					15	
Gly	Asn	Phe	Ser	Lys	Glu	Asp	Ala	Val	Leu	Gly	Asn	Gln	Phe	Pro	Ile
			20					25					30		
Leu	Lys	Lys	Ser	Asn	Asn	Leu	Ser	Ile	Ala	Arg	Pro	Pro	Ser	Ile	Glu
		35					40					45			
Ser	Phe	Ser	Ala	Ser	Val	Glu	Lys	Ile	Phe	Arg	Glu	Trp	Asn	Glu	Ser
	50					55					60				
Gly	Gly	Glu	Lys	Ile	Phe	Asp	Ile	Ser	Gln	Asn	Glu	Glu	Glu	Trp	Met
65					70					75					80
Asp	Ile	Ile	Ser	Leu	Val	Glu	Ser	Val	Tyr	Glu	Pro	Val	Phe	Ser	Lys
				85					90					95	
Ser	Leu	Lys	Pro	Asp	Lys	Leu	Ala	Asp	Lys	Thr	Cys	Leu	Thr	Ala	Ala
			100					105					110		
Ala	Phe	Ala	Ala	Ser	Ala	Val	Asp	Glu	Lys	Leu	Thr	Ile	Leu	Ser	Gly
		115					120					125			
Ser	Asp	Gly	Ser	Val	Leu	Gln	Arg	Thr	Thr	Lys	Val	Met	Lys	Lys	Asp
	130					135					140				
Pro	Lys	Lys	Ile	Ala	Glu	Ser	Leu	Leu	Asn	Asn	Glu	Lys	Trp	Thr	Ser
145					150					155					160
Ile	Leu	Leu	Asp	Arg	Leu	Lys	Thr	Ala	Lys	Lys	Leu	Leu	Ser	Arg	Arg
			165						170					175	
Gly	Ala	Leu	Lys	Ser	Ala	Glu	Arg	Val	Glu	Val	Leu	His	Arg	Leu	Asn
			180					185					190		
Lys	Leu	Lys	Glu	Ala	Pro	Leu	Pro	His	His	Pro	Ser	Leu	Phe	Asp	Asn
		195					200					205			
Phe	Ser	Gly	Gly	Lys	Thr	Ser	Ala	Val	Ser	Ala	Gly	Thr	Val	Ile	Ala
	210					215					220				
Ser	Asp	Met	His	Phe	Lys	Leu	Val	Glu	His	Ile	Phe	Lys	Val	Ser	Phe
225					230					235					240
Arg	Lys	Trp	Gly	Pro	Cys	Gly	Asp	Lys	Thr	Glu	Ser	Gly	Glu	Glu	Glu
				245					250					255	
Asp	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Lys	Lys	His	Ser	Ile	Ser	Arg	Phe
			260					265					270		
Val	Leu	Gln	Phe	Met	Asn	Gly	His	Asn	Gly	Gln	His	Tyr	His	Arg	Pro
		275					280					285			
Glu	Ser	Ala	Ser	Val	Tyr	Phe	Cys	Asp	Tyr	Tyr	Asp	Tyr	Leu	Ala	Tyr
	290					295					300				
Arg	Asn	Leu	Pro	Asn	Glu	Tyr	Lys	Leu	Ser	Ser	Met	His	Pro	Gly	Thr
305					310					315					320
Phe	Asn	Met	Glu	Asp	Leu	Pro	Phe	Arg	Pro	Phe	Ala	Val	Pro	Ser	Thr
				325					330					335	
Tyr	Lys	Thr	Glu	Leu	Glu	Tyr	Lys	Arg	Phe	Val	Gln	Ser	Thr	Asn	Leu
			340					345					350		
Pro	Gln	Leu	Ser	Phe	Asp	Tyr	Gly	Glu	Phe	Leu	Cys	Tyr	Cys	Ile	Phe
		355					360					365			
Gly	Ala	Asp	Trp	Tyr	Lys	His	Leu	Gly	Asp	Val	Val	Asp	Ser	Leu	Glu
	370					375					380				
Asn	Ser	Ser	Met	Ile	Ser	Phe	Asp	Ser	Gln	Thr	Leu	Ser	Gly	Val	Tyr
385					390					395					400

Lys Asn Thr Ala Asn Tyr Lys Arg Leu Gly Lys Lys Arg Asn Gly Ile  
 405 410 415  
 Ala Asp Leu Ala Val Arg Ser Met Ala Glu Phe Ile Arg Thr Glu Ala  
 420 425 430  
 His Lys Ala Leu Thr Ala Glu Glu Met Glu Glu Glu Glu Glu Glu  
 435 440 445  
 Glu Ala Glu Glu Glu Ala Met Asp Gln Glu Pro Ala Glu Val Asp Phe  
 450 455 460  
 Leu Ser Val Pro His Leu Arg Arg Lys Ile Arg Gln Ala Val Ser Val  
 465 470 475 480  
 Leu Asn Asn Phe Val Glu Asn Asp Leu Ser Ile Leu Val Ser Asn Phe  
 485 490 495  
 Lys Asn Val Leu Thr Asp Asp Thr Val Ser Gly Thr Asp Thr Asp Asn  
 500 505 510  
 Phe Gly Ser Ser Gly Glu Phe Glu Ala Leu Ser Ser His Leu Phe Leu  
 515 520 525  
 Ser Arg Ile Leu Asp Glu Val His Ile Leu Arg Asn Thr Asp Ile Gln  
 530 535 540  
 Arg Thr Leu Phe Ser Thr His Val Ser Leu Ser Asp Lys Ser Pro Pro  
 545 550 555 560  
 Ser Arg Val Arg Gly Ser Asn Val Asn Phe Asn Asn Asn Ala Gly Asn  
 565 570 575  
 Ile Ser Ser Leu Gln Thr Tyr Gly Gly Ile Glu Glu Leu Pro Glu Asn  
 580 585 590  
 Val Leu Val Gly Leu Ser Gly Gly Phe Glu Asp Thr Asp Met Tyr Ser  
 595 600 605  
 Gly Glu Asp Val Val Val Val Trp Asp Gly Cys Asp Gly Gly Lys Val  
 610 615 620  
 Leu Ser Val Thr Phe Asn Cys Gly Asp Asn Phe Ile Gln Leu His Glu  
 625 630 635 640  
 Lys Thr Ala Glu Thr Phe Lys Asp Asp Thr Asp Leu Val Glu Arg Ile  
 645 650 655  
 Arg Asp Val Leu Gln Thr Ala Ser Lys Thr Gly Asn Leu Asn Lys Lys  
 660 665 670  
 Ala Tyr Ser Arg Lys Asn Ile Tyr Ala Val Leu Arg Glu Asn Gly Ile  
 675 680 685  
 Glu Arg Pro Gly Asp Asp Phe Thr Glu Lys Gly Ile Ala Leu Lys Asp  
 690 695 700  
 Lys Thr Asn Gln Pro Pro Pro Pro Ala Arg Ser Ala Lys Ile Thr Val  
 705 710 715 720  
 Glu Gly Val Lys Gly Phe Phe Ser Gly Phe Arg Asp Ile Thr Arg Ala  
 725 730 735  
 Leu Thr Thr Tyr Ser Ala Glu Thr Phe Arg Asp Leu Gly Gln Gly Ile  
 740 745 750  
 Val Lys Glu Thr Glu Gly Leu Thr Ala Ala Thr Val Ala Glu Thr Ser  
 755 760 765  
 Phe Ser Glu Gly Leu Ala Glu Ser Leu Arg Ser Asp Ala Asn Leu Gly  
 770 775 780  
 Leu Glu Phe Ser Glu Asp Ala Lys Thr Val Val Phe Lys Asn Asp Thr  
 785 790 795 800  
 Ser Arg Ser Leu Leu Glu Glu Thr Arg Ala Leu Arg Ala Asn Asn Thr  
 805 810 815  
 Ser Phe Ser Ser Phe Ala Arg Asp Met Gly Val Gln Val Ser Ala Asp  
 820 825 830  
 Leu Asp Ala Glu Phe Ala Ala Glu Met Arg Glu Thr Tyr Pro Asp Ala  
 835 840 845  
 Ala Leu Glu Gln Asn Leu Lys Asp Leu Asp Lys Phe Glu Glu Thr Ile  
 850 855 860  
 Pro Glu Ser Gln Val Lys Lys Leu Lys Lys Ile Asp Ser Tyr Leu Thr  
 865 870 875 880  
 Glu Asn Pro Glu Arg Ala Gly Lys Glu Ile Asn Asp Thr Glu Leu Ser

885 890 895  
 Lys Ala Thr Asp Ser Val Leu Gly Lys Lys Leu Gly Asn Ala Val Thr  
 900 905 910  
 Val Leu Met Asn Asn Phe Gly Lys Val Thr Ile Val Val Gly Ala Ser  
 915 920 925  
 Val Val Ala Gly Phe Leu Gly Pro Ala Ala Val Val His Ala Ser Arg  
 930 935 940  
 Gly Ala His Leu Asn Val Val Asp His Thr Ser Pro Lys Gly Val Ile  
 945 950 955 960  
 Ser Tyr Lys Ile Val Asp Phe Ser Cys Ala Asp Arg Asn Thr Gly Trp  
 965 970 975  
 Ala Lys Pro Thr Lys His Pro Phe Arg Glu Glu Ile Asp His Val Ile  
 980 985 990  
 Ala Leu Asp Ala Ser Phe Leu Thr Glu Asn Gly Ala Tyr Val Phe Pro  
 995 1000 1005  
 Glu Asp Gly Gly Pro Lys Ser Lys Tyr Lys Ala Tyr Ala Pro Ile Cys  
 1010 1015 1020  
 Gly Thr Lys Asp Ala Ala Gln Gly Glu Cys Gly Ser Trp Ala Thr Phe  
 1025 1030 1035 1040  
 Asp Asp Pro His Ser Val Leu Pro Trp Val Ala Ser Met Lys Asp Leu  
 1045 1050 1055  
 Pro Lys Gly Gln Ser Leu Ser Cys Asp Lys Gly Met Ser Thr Leu Lys  
 1060 1065 1070  
 Ala Val Ser Ser Val Leu Leu Ser Ile Gly Lys Asp Val Ala Glu Ala  
 1075 1080 1085  
 Ile Phe Glu Val Ala Glu Asp Ala Val Val Gly Leu Ala Ser Lys Ala  
 1090 1095 1100  
 Ile Ser Ala Val Ile Asn Asn Pro Leu Phe Ile Phe Gly Val Pro Leu  
 1105 1110 1115 1120  
 Gly Phe Gly Ile Ala Ala Thr Arg Leu Asn Pro Ser Asn Trp Lys Thr  
 1125 1130 1135  
 Gly Leu Ile Val Phe Ser Ile Leu Leu Val Val Ile Leu Ile Val Arg  
 1140 1145 1150  
 Phe Phe Ala Gly Ser Gly Pro Leu Thr Leu Asn Trp Phe Gly Ala Lys  
 1155 1160 1165  
 Asn Ser Ala Lys Arg Lys Gln Thr Glu Gln Phe Glu Asp Gly Gly Gly  
 1170 1175 1180  
 Asn Arg Ser Lys Ile Val Leu Ala Glu Lys Asp Asn Ala Asn Ser Lys  
 1185 1190 1195 1200  
 Leu Gln Ser Arg Arg Asn Glu Thr Gly Pro Met Arg Leu Glu Glu Leu  
 1205 1210 1215  
 Pro Gly His Glu Asp Leu Arg Pro Val Phe Phe Pro Ala Thr Asn  
 1220 1225 1230  
 Tyr Ser Lys Ser Ala Lys Ile Leu Gly Tyr Lys Ser Lys Pro Phe Asn  
 1235 1240 1245  
 Asp Phe Tyr Thr Lys Ile Ile Asn Thr Asp Ile Ile Lys Met Asp Arg  
 1250 1255 1260

<210> 212  
 <211> 3744  
 <212> DNA  
 <213> SHRIMP

<400> 212  
 atgggcgcgc ctaccaacgc tgattttaca cgcacggtat caggagtagc ttcttctctc 60  
 tatcttggtta accctggagc tccttccgat agagaaaagt tagttttggc ctcttcttat 120  
 tctgactctt ttgtgtacaa ctacaaggat gcagtgggtga ccgctgaggc tcccaagtgg 180  
 tgtcccttta acgagccagc tcttcatgag cacatcatga acagacttga aaaagctggg 240  
 ctaattaaca gatctcgttt tgtgtgtaac cctgttaa at cggctggaga agtatgcgga 300  
 tttcgctatt ctggaggaag tactccccag aacttaattt tcccgattgg agcatcagag 360

ataaattggt	tggtcatact	tagaaatgct	gctcgttttg	gtacagtggc	agcatcggcc	420
aaagacgcc	tcgaacgc	tcctgatcta	agagaagggt	gtacaagtaa	acatgtggca	480
aagaatgcaa	tgaggagact	tcgtgtatgg	cgagccttta	actggatagc	ggaagcctcc	540
agatcagcgg	gcatgatctg	ctacgaaccc	ttttcggttg	gctgtgctct	atacagatcac	600
gaggtaaagaa	gaaggcaact	aaaagggaagt	taggagcgaa	atgtactatt	tcttggtgac	660
catttctgta	aagaatcaac	tcttctggct	gatatttcaa	gagggggaag	aagttctgat	720
ttctggacga	tcgttgaagc	tgtcatccgg	tacaagaata	gacatgctcg	aacaatcagt	780
aatgaaacta	atgccatccc	tgaagactct	tctataaact	tggagtggga	ggatgttcta	840
gtaaaaaacc	aacgcgcac	aatcaagga	gatgattcta	cgttagaaaa	gaccttagaa	900
gctgccatca	aagaacatga	aagcatagga	gaaaaaagga	aaaaacatat	cctagagtgt	960
attaaacat	gtctaacaga	ggagcaacga	gaaatgattt	ttaagggaag	tggaggaaaa	1020
gggaatttgt	ccccggccca	tctaacaat	ctggccgatg	cgatactggc	taataatgcc	1080
aaagcaggca	tatgggttat	tatgcagagt	cttttaaagc	aaatcaattt	ctctatactc	1140
catttaatat	ggtacgaagc	tcaacgactt	ttaatgttca	agttgtatat	gcctgcatta	1200
cttgctcttt	ttatatctca	gaggggaatt	ggtgacgtgt	ttctgaatgg	tgtgtttaac	1260
ctagaagtga	tgaagaaaag	agctgcaaat	gcaaaaaata	gagacatggt	ttctcgagac	1320
gcctacaaga	acaatactaa	cgagtcta	aatgtgggca	attattccca	attcgatata	1380
gtacaatttt	atgggtcaaat	gtctgattgt	aatgcaattc	ctctatctat	taatatggga	1440
gtacctctcc	ataaatcgag	gatgaacatg	caagacattg	aaaagaccat	acaacacata	1500
gtggatatgg	gtttggcaaa	gatacgtgta	gcggaagaaa	atagagttagc	tagacgtctt	1560
ttacatagaa	gaaaaatgca	agaaaaggca	gccagggagc	gagcagctgc	gagacagaga	1620
agactaagag	gtgaagaaga	agaagaagaa	gaggaggagg	aggaagaggg	agaagaaatg	1680
gaagaagagg	aagaagaagc	tggaaacatca	ggtgttaatg	gttctgggta	tgatcaagaa	1740
gaagaagatg	atggtgagga	ggaggaggag	gaagagggaag	aagaggatga	agaagatagt	1800
gagggtgaag	atatgaatgg	ggaaaaaac	tcaagaaaa	gtaaaaaac	tgtaataact	1860
tcctccactc	aacaacctcc	ccaaaagcgt	cagcgaggta	aaaatgcgcc	catttcaacc	1920
aagggaaaaa	ggttaaaggga	aagagataat	attggaggat	ttttattagc	cacaattcaa	1980
aacgatgacc	gccaaagtga	tgtggagagc	atacaaaaa	ttttgacggc	caggcaaaag	2040
aaatatgtta	aaggtggaaa	atgtggagat	aatggactac	cagaattggt	aattgaaaag	2100
gttacaatc	tactagattc	agtgttccat	ttcagaaaag	ggtctattct	taacagcata	2160
cacgcaaat	ggcgatcaga	aactggcgta	tacaccacaa	aggcaaaact	tatttgtgat	2220
tactatgaaa	gaaacgtatc	aaaagacagt	agtaataata	cacccattc	atctgaatgt	2280
attgagcggg	caagagaaa	ggatgcttcc	tgtgccgaat	ctaacaacg	cccttgctct	2340
gtagactcta	acaacctga	agatgtggaa	caacgtatgc	gggaattgat	catggacct	2400
ccttcattat	caggggtaga	agattctctc	gctatcgaaa	gagtacttca	aaatgagata	2460
ttattcacaa	gtcttgctac	caatccaatt	ttcaatgccg	tgttaggtgc	tgaaaaagga	2520
gatcttgga	gatttatcgt	attaaataac	attgtaaaa	ttatgaatat	gactattgct	2580
tgtctagtgg	acggagatat	gcctatgctt	ctagactcaa	gaggcaagac	aaaaaacctt	2640
ctagaaaagg	gtacagtga	aaacacgaga	aaatttttca	aaccaaatat	gactgcagca	2700
gaattgaacg	tggccactgc	tcaatctgca	ggacaccaat	acatgaacgc	tggtcattgt	2760
ccagaacctg	gaatcaagca	aatgttactc	cctgattgta	taatgaaact	gaaatccatc	2820
gccatggaaa	agggtcgagg	aggagatcg	gcccttcaca	gacagaaatg	tgaccatgcc	2880
ttttgcaaga	tgctgaagtg	tttattcttc	aatattgacc	cttcaaattg	tgcagatata	2940
tttattgacc	ctgcgtcacg	tgccacctta	tttaggcttg	atgatctttg	tagggatagg	3000
aaaaaataca	agaacattga	ctgggtaaa	gatttactag	accctgtaat	gaagggaaca	3060
aataaatggg	tgggaactgg	agaatacact	aacattggac	gtgactcaaa	cgtggccgcc	3120
cctgttgatt	tttacacaat	tttgaaatc	acaaatgatt	atgatgggtg	aataagttgt	3180
ccttcacgga	aaccaaatga	cgtgtattac	agtactattg	aaagagcgga	tgacttactc	3240
accgagagca	gggacgcttc	atgtgaaagc	taccgcccga	ccttggttga	cgccagagcg	3300
gtgttggaag	tcaacggaga	cggacgtgtc	ccttaccctt	ccagtgaac	cgtagaagac	3360
ttgggaggag	aagaagaaga	agaggaaata	acaggcggtta	ttgacgacag	tactgaaata	3420
gaagacgttc	aagtacagga	cagtaatcta	ttcgatgtgg	aattattcga	tatccctgaa	3480
atagaacaac	atcaacaggg	tggagaagaa	gaggaaactc	cttcagcaat	atctgaagtg	3540
tttgcttcat	taccagctga	taacgattcc	tcctcaccgg	cgcatattcc	ctcatttggg	3600
aattctgagg	agggggaaaa	aagtccagaa	ccgtacaata	tttttgattc	tgccctcgac	3660
caattatttg	atttaataga	tagcgatgga	agaaacaata	accctaaaag	agtcgactgg	3720
aacagtgtca	ccattcaaga	atga				3744

<210> 213  
 <211> 1245  
 <212> PR

<213> SHRIMP

<400> 213

Met 1	Gly 2	Ala 3	Pro 4	Thr 5	Asn 6	Ala 7	Asp 8	Phe 9	Thr 10	Arg 11	Thr 12	Val 13	Ser 14	Gly 15	Val 16
Ala	Ser	Ser	Leu 20	Tyr	Leu	Val	Asn 25	Pro	Gly	Ala	Pro	Ser	Asp 30	Arg	Glu
Lys	Leu 35	Val	Leu	Ala	Ser	Ser	Tyr 40	Ser	Asp	Ser	Phe	Val 45	Tyr	Asn	Tyr
Lys	Asp 50	Ala	Val	Val	Thr	Ala 55	Glu	Ala	Pro	Lys	Trp 60	Cys	Pro	Phe	Asn
Glu 65	Pro	Ala	Leu	His 70	Glu	His	Ile	Met	Asn 75	Arg	Leu	Glu	Lys	Ala	Gly 80
Leu	Ile	Asn	Arg	Ser 85	Ser	Arg	Phe	Val	Cys 90	Asn	Pro	Val	Lys	Ser	Ala 95
Glu	Val	Cys	Gly 100	Phe	Arg	Tyr	Ser	Gly 105	Gly	Ser	Thr	Pro	Gln 110	Asn	Leu
Ile	Phe	Pro	Ile 115	Gly	Ala	Ser	Glu 120	Ile	Asn	Trp	Leu	Val 125	Ile	Leu	Arg
Asn	Ala 130	Ala	Arg	Phe	Gly	Thr 135	Val	Ala	Ala	Ser	Ala 140	Lys	Asp	Ala	Ile
Glu 145	Arg	Ile	Pro	Asp 150	Leu	Arg	Glu	Gly	Gly	Thr 155	Ser	Lys	His	Val	Ala 160
Lys	Asn	Ala	Met 165	Arg	Arg	Leu	Arg	Val	Trp 170	Arg	Ala	Phe	Asn	Trp 175	Ile
Ala	Glu	Ala	Ser 180	Arg	Ser	Ala	Gly	Met 185	Ile	Arg	Tyr	Glu	Pro 190	Phe	Ser
Val	Gly	Cys 195	Ala	Leu	Tyr	Asp	His 200	Glu	Val	Arg	Arg	Arg 205	Gln	Leu	Lys
Gly	Ser 210	Tyr	Glu	Arg	Asn	Val 215	Leu	Phe	Leu	Gly	Asp 220	His	Phe	Cys	Lys
Glu 225	Ser	Thr	Leu	Leu 230	Ala	Asp	Ile	Ser	Arg	Gly 235	Gly	Arg	Ser	Ser	Asp 240
Phe	Trp	Thr	Ile 245	Val	Glu	Ala	Val	Ile	Arg 250	Tyr	Lys	Asn	Arg	His 255	Ala
Arg	Thr	Ile 260	Ser	Asn	Glu	Thr	Asn 265	Ala	Ile	Pro	Glu	Asp 270	Ser	Ser	Ile
Asn	Leu	Glu 275	Trp	Glu	Asp	Val	Leu 280	Val	Lys	Asn	Gln	Arg 285	Asp	Thr	Asn
Gln	Gly 290	Asp	Asp	Ser	Thr	Leu 295	Glu	Lys	Thr	Leu	Glu 300	Ala	Ala	Ile	Lys
Glu 305	His	Glu	Ser	Ile 310	Gly	Glu	Lys	Arg	Lys	Lys 315	His	Ile	Phe	Ile	Lys 320
Thr	Cys	Leu	Thr 325	Glu	Gln	Arg	Glu	Met 330	Ile	Phe	Lys	Gly	Val 335	Gly	Gly
Gly	Lys	Gly 340	Asn	Leu	Ser	Pro	Ala	His 345	Leu	Thr	Asn	Leu	Ala 350	Asp	Ala
Ile	Leu 355	Ala	Asn	Asn	Ala	Lys	Ala 360	Gly	Ile	Trp	Val	Ile 365	Met	Gln	Ser
Leu	Leu 370	Lys	Gln	Ile	Asn	Phe 375	Ser	Ile	Leu	His	Leu 380	Ile	Gly	Tyr	Glu
Ala 385	Gln	Arg	Leu	Leu 390	Met	Phe	Lys	Leu	Tyr	Met 395	Pro	Ala	Leu	Leu	Ala 400
Leu	Phe	Ile	Ser 405	Gln	Arg	Gly	Ile	Gly	Asp 410	Val	Phe	Leu	Asn	Gly 415	Val
Phe	Asn	Leu 420	Glu	Val	Met	Lys	Glu	Arg 425	Ala	Ala	Asn	Ala 430	Lys	Ile	Arg
Asp	Met 435	Val	Ser	Arg	Asp	Ala	Tyr 440	Lys	Asn	Asn	Thr	Asn 445	Glu	Ser	Asn
Asn	Leu 450	Gly	Asn	Tyr	Ser	Gln 455	Phe	Asp	Ile	Ala	Thr 460	Ile	Tyr	Gly	Gln

WO 01/38351

PCT/US00/28888

300

```

Met Ser Asp Cys Asn Ala Ile Pro Leu Ser Ile Asn Ile Gly Val Pro
465                               470                               475                               480
Leu His Lys Ser Arg Met Asn Met Gln Asp Ile Glu Lys Thr Ile Gln
                               485                               490                               495
His Ile Val Asp Met Gly Leu Ala Lys Ile Arg Val Ala Glu Glu Asn
                               500                               505                               510
Arg Val Ala Arg Arg Leu Leu His Arg Arg Lys Met Gln Glu Lys Ala
                               515                               520                               525
Ala Arg Glu Arg Ala Ala Ala Arg Gln Arg Arg Leu Arg Gly Glu Glu
                               530                               535                               540
Glu Glu Glu Glu Glu Glu Glu Glu Glu Gly Glu Glu Met Glu Glu
545                               550                               555                               560
Glu Glu Glu Glu Ala Gly Thr Ser Gly Val Asn Gly Ser Gly Tyr Asp
                               565                               570                               575
Gln Glu Glu Glu Asp Asp Gly Glu Glu Glu Glu Glu Glu Glu
                               580                               585                               590
Glu Asp Glu Glu Asp Ser Glu Gly Glu Asp Met Asn Gly Glu Asn Asn
                               595                               600                               605
Ser Arg Lys Arg Lys Asn Thr Gly Asn Thr Ser Ser Thr Gln Gln Pro
610                               615                               620
Pro Gln Lys Arg Gln Arg Gly Lys Asn Ala Pro Ile Ser Thr Lys Gly
625                               630                               635                               640
Lys Arg Leu Lys Glu Arg Asp Asn Ile Gly Gly Phe Leu Leu Ala Thr
                               645                               650                               655
Ile Gln Asn Asp Asp Arg Gln Val Asn Val Glu Ser Ile Gln Lys Leu
660                               665                               670
Leu Thr Ala Arg Gln Arg Lys Tyr Val Lys Gly Gly Lys Cys Gly Asp
675                               680                               685
Asn Gly Leu Pro Glu Leu Leu Ile Glu Lys Val Thr Asn Leu Leu Asp
690                               695                               700
Ser Val Phe His Phe Arg Lys Gly Ser Ile Leu Asn Ser Ile His Ala
705                               710                               715                               720
Asn Arg Arg Ser Glu Thr Gly Val Tyr Thr Thr Lys Ala Asn Cys Ile
725                               730                               735
Cys Asp Tyr Tyr Glu Arg Asn Val Ser Lys Asp Ser Ser Asn Asn Thr
740                               745                               750
Pro His Ser Ser Glu Cys Ile Glu Arg Ala Arg Glu Arg Asp Ala Ser
755                               760                               765
Cys Ala Glu Ser Asn Lys Arg Pro Cys Pro Val Asp Ser Asn Asn Pro
770                               775                               780
Glu Asp Val Glu Gln Arg Met Arg Glu Leu Ile Met Asp Pro Pro Ser
785                               790                               795                               800
Leu Ser Gly Val Glu Asp Ser Leu Ala Ile Glu Arg Val Leu Gln Asn
805                               810                               815
Glu Ile Leu Phe Thr Ser Leu Val Thr Asn Pro Ile Phe Asn Ala Val
820                               825                               830
Leu Gly Ala Glu Lys Gly Asp Leu Gly Arg Phe Ile Val Leu Asn Asn
835                               840                               845
Ile Val Lys Phe Met Asn Met Thr Ile Ala Cys Leu Val Asp Gly Asp
850                               855                               860
Met Pro Met Leu Leu Asp Ser Arg Gly Lys Thr Lys Asn Leu Leu Glu
865                               870                               875                               880
Lys Gly Thr Val Lys Asn Thr Arg Lys Phe Phe Lys Pro Asn Met Thr
885                               890                               895
Ala Ala Glu Leu Asn Val Ala Thr Ala Gln Ser Ala Gly His Gln Tyr
900                               905                               910
Met Asn Ala Gly His Cys Pro Glu Pro Gly Ile Lys Gln Met Leu Leu
915                               920                               925
Pro Asp Cys Ile Met Lys Leu Lys Ser Ile Ala Met Glu Lys Gly Arg
930                               935                               940
Gly Gly Arg Ser Ala Leu His Arg Gln Lys Cys Asp His Ala Phe Cys

```



WO 01/38351

301

PCT/US00/28888

945 950 955 960  
Lys Met Leu Lys Cys Leu Phe Phe Asn Ile Asp Pro Ser Asn Ala Ala  
965 970 975  
Asp Thr Phe Ile Asp Pro Ala Ser Arg Ala Thr Leu Phe Arg Leu Asp  
980 985 990  
Asp Leu Cys Arg Asp Arg Lys Lys Tyr Lys Asn Ile Asp Trp Val Lys  
995 1000 1005  
Asp Leu Leu Asp Pro Val Met Lys Gly Thr Asn Lys Trp Val Gly Thr  
1010 1015 1020  
Gly Glu Tyr Thr Asn Ile Gly Arg Asp Ser Asn Val Ala Ala Pro Val  
1025 1030 1035 1040  
Asp Phe Tyr Thr Ile Leu Lys Tyr Thr Met Ile Asp Asp Gly Val Ile  
1045 1050 1055  
Ser Val Pro Ser Arg Lys Pro Asn Asp Val Tyr Tyr Ser Thr Ile Glu  
1060 1065 1070  
Arg Ala Asp Asp Leu Leu Thr Glu Ser Arg Asp Ala Ser Cys Glu Ser  
1075 1080 1085  
Tyr Arg Pro Thr Leu Phe Asp Ala Arg Ala Val Leu Glu Val Asn Gly  
1090 1095 1100  
Asp Gly Arg Val Pro Tyr Pro Ser Ser Glu Thr Val Glu Asp Leu Gly  
1105 1110 1115 1120  
Gly Glu Glu Glu Glu Glu Glu Ile Thr Gly Val Ile Asp Asp Ser Thr  
1125 1130 1135  
Glu Ile Glu Asp Val Gln Val Gln Asp Ser Asn Leu Phe Asp Val Glu  
1140 1145 1150  
Leu Phe Asp Ile Pro Glu Ile Glu Gln His Gln Gln Gly Gly Glu Glu  
1155 1160 1165  
Glu Glu Leu Pro Ser Ala Ile Ser Glu Val Phe Ala Ser Leu Pro Ala  
1170 1175 1180  
Asp Asn Asp Ser Ser Ser Pro Ala His Ile Pro Ser Phe Gly Asn Ser  
1185 1190 1195 1200  
Glu Glu Gly Glu Lys Ser Pro Glu Pro Tyr Asn Ile Phe Asp Ser Ala  
1205 1210 1215  
Leu Asp Gln Leu Leu Asp Leu Ile Asp Ser Asp Gly Arg Asn Asn Asn  
1220 1225 1230  
Pro Lys Arg Val Asp Trp Asn Ser Val Thr Ile Gln Glu  
1235 1240 1245

<210> 214  
<211> 2946  
<212> DNA  
<213> SHRIMP

<400> 214  
atggaagaaa caataaccct aaaagagtcg actggaacag tgaccaccatt caagaatgaa 60  
aatattacca gaatagcgtc taattacgtc cgagccttta ctgacacgtg gtctcatttg 120  
gtcaacattt ctggagctcc tctaactgct gagaagaatc caagcgctat cccagctaac 180  
gaactgaata gatactggac caaaactaac gtgttatgca acccactctt taaattggag 240  
gaccacataa cgaggatga agatactggc acaataacac taaaattcaa aatgtatata 300  
gatgataaaa atggactata tcagctctgcc gttttaatgc tggctctcga ttcgttcgtt 360  
tcgcttgcat ctttttccca tggagctgat ttggtttcca ataaaagtga aaacaaattc 420  
tgcgtaaaaa ttccccacga cactcgcgct gaatctttac tgaataatgt tggattcccc 480  
gctggattaa gtgggccttt taagagatgg agcattaact acaaggctgc aaacttgagc 540  
ggtaaaagcg gtatagatgg cctttcaggt tccatgttga cagtactaaa aaataacacc 600  
aataaaagag caactgatat ttacattttg gtgaataatg tttctgcctc tgcacaacaa 660  
cttgacgatt ctgaaatgtc tcgtactttt aaccacaaaa agaaagttag agtttgttat 720  
gatataaatg tgtccagttc aaggcaagtg aaccagcgta atttactcca ccatcagaat 780  
ataataggac aacatctgat cgaattttag accaagcaac ttgaacgcgc acaaaaataa 840  
aaagtcaagg aagaggaaaa tgggtgagcat gaagaaatga caagtgagga ggaagaagag 900  
gaagatgaat atgaagaagg aggttgcccta tcagatattg acgaggaaga tttctatgag 960

```

gatgggttacg atgaagaaga gggcgatgac aatagaacta gaaagaagaa gaaaatggaa 1020
gaagatgaag aggatgaaga agaagaatat gacgatgaag aagatgaaga agaggcagaa 1080
acttgtggtg ctaatggtgt tattgattgt gaagacgatg caatcatttt ccccaatgga 1140
caaaattcaa aaaggaagaa aaatggtaaa aaaacaaaca ttaaaaagcg gtcacggagg 1200
aagggggagtg gctctgctaa cactttatcc tttgtggaaa aatacgttgg aaattgtaag 1260
agcctaggta taaagccagt aggggtgtccg ccccatccca ctgaatttac ttccctattt 1320
atgaagggga gtgaagctga cagctgttat aatacttgtc agtccaccag aggggctagc 1380
cgtataaggt cactactcaa taaatactct gttaaagatt tgatgcaggt aaacagccct 1440
tcgagttgga aatgggctaa ccctcccgac cgacggttcg tgctgtttga taagaaaact 1500
aaggaagaag ttgagggttaa gtttgaaatt gaatgtgaaa aatccgagta ttttgatgtc 1560
gtatctgaac tccctagtaa tattaagta tggttaaaag agacggcaaa aataataaaa 1620
catttggtc tcgattgaaga ctttcttcca gctatgggtg ctgctacccc aaaaattccc 1680
ctcaatttga ttaaaactat gacgagcatt ttctctgtta gagatattgt tggatttaaa 1740
ataccagaag aagtgtctcag ttttattcct atagaatgga agacatctat ttctgcaatg 1800
gggtcctct ctgtacaatt tgatcgata atagaagtga tagatttaat gataactaat 1860
ggcgcccttg cgagtcctg cttgaacaac gcattcttct tggaaagagg agtgggtgcc 1920
agagatggga gtaacacgtg gctccacacg gaccttgtgc aactctccac ctccatattt 1980
agaagtattc gcaacagagg agtgaatatt ggcggtaca acaacactgg tagcaattct 2040
tctagttctt cttgtggagg gaataagggc gattatggag tacgttgtgg attgagtata 2100
agcaagcgtg gtataaccct gaaaccacca cctgcagcga tgactaattc ttcttccccg 2160
tcctcctcgg ccatgatctc attgcctcag cccacgcgcc agagcataga tctttcgata 2220
acgacaatca tccaagattt ctcaagaagt tctgggaaat tgaggcttaa tggattacag 2280
aaaaacatgt ctgacaagag caaagacgtg tttaatgatg caatatacga ctctggcgca 2340
ttcaaggcgc tcctaacatg cacagtcaac gataaaaagta gacgtaaaag gaaaagaagg 2400
actttatttg catctggaga ggggtgtggt cgaagaaacc tgatgggtgag tcagggcaat 2460
gacgtcaatg atgcccacca gttccaggaa gaatgcggaa taaaatttg gggcggggct 2520
tctagggtgt ataaaagagc ccagcgccga gggttcggcag tcagttccag aagaagagta 2580
aggaacaaac cccagtttac tatagcagtc tctgacgaag acgacgactg cgaagaagaa 2640
ggcgactttt ctcccgagtt aaatccaacg cactctcaac tacttctatt ccaacaacgg 2700
caacaagata gctgcacaga agacgacgac gtttttagtgt ctgtagaaga atataacaac 2760
agagtaagcg gttcttccac cacagccgga gacagagttc ttgcaaagga tcttctctct 2820
actgtatctc cgaacgaaaa gaggaactct cgcgcctcgc cgcactcac catatcccg 2880
cactctcttt tcaacgctct atctgcaaaa acaaagttgg gagaaaatgg acgtttcttc 2940
ctataa 2946

```

&lt;210&gt; 215

&lt;211&gt; 975

&lt;212&gt; PRT

&lt;213&gt; SHRIMP

&lt;400&gt; 215

```

Met Glu Glu Thr Ile Thr Leu Lys Glu Ser Thr Gly Thr Val Ser Pro
1          5          10          15
Phe Lys Asn Glu Asn Ile Thr Arg Ile Asn Tyr Val Arg Ala Phe Thr
20          25          30
Asp Thr Trp Ser His Leu Val Asn Ile Ser Gly Ala Pro Leu Thr Ala
35          40          45
Glu Lys Asn Pro Ser Ala Ile Pro Ala Asn Glu Leu Asn Arg Tyr Trp
50          55          60
Thr Lys Thr Asn Val Leu Cys Asn Pro Leu Phe Lys Leu Glu Asp His
65          70          75          80
Ile Thr Arg Asp Glu Asp Thr Gly Thr Ile Thr Leu Lys Phe Lys Met
85          90          95
Tyr Ile Asp Asp Lys Asn Gly Gln Ser Ala Val Leu Met Leu Ala Leu
100          105          110
Asp Ser Phe Val Ser Leu Ala Ser Phe Ser His Gly Ala Asp Leu Val
115          120          125
Ser Asn Lys Ser Glu Asn Lys Phe Cys Val Lys Ile Pro His Asp Thr
130          135          140
Arg Ala Glu Ser Leu Leu Asn Asn Val Gly Phe Pro Ala Gly Leu Ser
145          150          155          160

```

Gly Pro Phe Lys Arg Trp Ser Ile Asn Tyr Lys Ala Ala Asn Leu Ser  
 165 170 175  
 Gly Lys Ser Gly Ile Asp Gly Leu Ser Gly Ser Met Leu Thr Val Leu  
 180 185 190  
 Lys Asn Asn Thr Asn Lys Arg Ala Thr Asp Ile Leu His Leu Val Asn  
 195 200 205  
 Asn Val Ser Ala Ser Ala Gln Gln Leu Asp Asp Ser Glu Met Ser Arg  
 210 215 220  
 Thr Phe Asn His Gln Lys Lys Val Gly Val Cys Tyr Asp Ile Asn Val  
 225 230 235 240  
 Ser Ser Ser Arg Gln Val Asn Gln Arg Asn Leu Leu His His Gln Asn  
 245 250 255  
 Ile Ile Gly Gln His Leu Ile Glu Phe Arg Thr Lys Gln Leu Glu Arg  
 260 265 270  
 Ala Gln Asn Lys Lys Val Lys Glu Glu Asn Gly Glu His Glu Glu  
 275 280 285  
 Met Thr Ser Glu Glu Glu Glu Glu Glu Asp Glu Tyr Glu Glu Gly Gly  
 290 295 300  
 Cys Leu Ser Asp Ile Asp Glu Glu Asp Phe Tyr Glu Asp Gly Tyr Asp  
 305 310 315 320  
 Glu Glu Glu Gly Asp Asn Arg Thr Arg Lys Lys Lys Met Glu  
 325 330 335  
 Glu Asp Glu Glu Asp Glu Glu Glu Glu Tyr Asp Asp Glu Glu Asp Glu  
 340 345 350  
 Glu Glu Ala Glu Thr Cys Gly Ala Asn Gly Val Ile Asp Cys Glu Asp  
 355 360 365  
 Asp Ala Ile Ile Phe Pro Asn Gly Gln Asn Ser Lys Arg Lys Lys Asn  
 370 375 380  
 Gly Lys Lys Thr Asn Ile Lys Lys Arg Ser Arg Arg Lys Gly Glu Cys  
 385 390 395 400  
 Ser Ala Asn Thr Leu Ser Phe Val Glu Lys Tyr Val Gly Asn Cys Lys  
 405 410 415  
 Ser Leu Gly Ile Lys Pro Val Gly Cys Pro Pro Pro Ser Thr Glu Phe  
 420 425 430  
 Thr Ser Leu Phe Met Lys Gly Ser Glu Ala Asp Ser Cys Tyr Asn Thr  
 435 440 445  
 Cys Gln Ser Thr Arg Gly Ala Ser Arg Ile Arg Ser Leu Leu Asn Lys  
 450 455 460  
 Tyr Ser Val Lys Asp Leu Met Gln Val Asn Ser Pro Ser Ser Trp Lys  
 465 470 475 480  
 Trp Ala Asn Pro Pro Asp Arg Arg Phe Val Leu Phe Asp Lys Lys Thr  
 485 490 495  
 Lys Glu Glu Val Glu Val Lys Phe Glu Ile Glu Cys Glu Lys Ser Glu  
 500 505 510  
 Tyr Phe Asp Val Val Ser Glu Leu Pro Ser Asn Ile Lys Val Trp Leu  
 515 520 525  
 Lys Glu Thr Ala Lys Ile Ile Lys His Leu Ala Leu Ile Glu Asp Phe  
 530 535 540  
 Leu Pro Ala Met Gly Ala Ala Thr Pro Lys Ile Pro Leu Asn Leu Ile  
 545 550 555 560  
 Lys Thr Met Thr Ser Ile Phe Ser Val Arg Asp Ile Val Gly Phe Lys  
 565 570 575  
 Ile Pro Glu Glu Val Leu Ser Phe Ile Pro Ile Glu Trp Lys Thr Ser  
 580 585 590  
 Ile Ser Ala Met Gly Leu Leu Ser Val Gln Phe Asp Arg Ile Ile Glu  
 595 600 605  
 Val Ile Asp Leu Met Ile Thr Asn Gly Ala Phe Ala Thr Ser Cys Leu  
 610 615 620  
 Asn Asn Ala Phe Phe Leu Glu Arg Gly Val Val Pro Arg Asp Gly Ser  
 625 630 635 640  
 Asn Thr Trp Leu His Thr Asp Leu Val Gln Leu Ser Thr Ser Ile Phe

Arg	Ser	Ile	Arg	Asn	Arg	Gly	Val	Asn	Ile	Gly	Gly	Asn	Asn	Asn	Thr
Gly	Ser	Asn	Ser	Ser	Ser	Ser	Ser	Cys	Gly	Gly	Asn	Lys	Gly	Asp	Tyr
Gly	Val	Arg	Cys	Gly	Leu	Ser	Ile	Ser	Lys	Arg	Gly	Ile	Thr	Leu	Lys
Pro	Pro	Pro	Ala	Ala	Met	Thr	Asn	Ser	Ser	Ser	Pro	Ser	Ser	Ser	Ala
Met	Ile	Ser	Leu	Pro	Gln	Pro	Thr	Arg	Gln	Ser	Ile	Asp	Leu	Ser	Ile
Thr	Thr	Ile	Ile	Gln	Asp	Phe	Ser	Glu	Val	Ser	Gly	Lys	Leu	Arg	Leu
Asn	Gly	Leu	Gln	Lys	Asn	Met	Ser	Asp	Lys	Ser	Lys	Asp	Val	Phe	Asn
Asp	Ala	Ile	Tyr	Asp	Ser	Gly	Ala	Phe	Lys	Ala	Leu	Leu	Thr	Cys	Thr
Val	Asn	Asp	Lys	Ser	Arg	Arg	Lys	Arg	Lys	Arg	Arg	Thr	Leu	Leu	Ala
Ser	Gly	Glu	Gly	Val	Val	Arg	Arg	Asn	Leu	Met	Val	Ser	Gln	Gly	Asn
Asp	Val	Asn	Asp	Ala	His	Gln	Phe	Gln	Glu	Glu	Cys	Gly	Ile	Lys	Ile
Gly	Gly	Gly	Ala	Ser	Arg	Val	Tyr	Lys	Arg	Ala	Gln	Arg	Arg	Gly	Ser
Ala	Val	Ser	Ser	Arg	Arg	Arg	Val	Arg	Asn	Lys	Pro	Gln	Phe	Thr	Ile
Ala	Val	Ser	Asp	Glu	Asp	Asp	Asp	Cys	Glu	Glu	Glu	Gly	Asp	Phe	Ser
Ser	Glu	Leu	Asn	Pro	Thr	His	Ser	Gln	Leu	Leu	Leu	Phe	Gln	Gln	Arg
Gln	Gln	Asp	Ser	Cys	Thr	Glu	Asp	Asp	Asp	Val	Leu	Val	Ser	Val	Glu
Glu	Tyr	Asn	Asn	Arg	Val	Ser	Gly	Ser	Ser	Thr	Thr	Ala	Gly	Asp	Arg
Val	Leu	Ala	Lys	Asp	Leu	Leu	Ser	Thr	Val	Ser	Pro	Asn	Glu	Lys	Arg
Asn	Ser	Ala	Ala	Ala	Leu	Thr	Ile	Ser	Arg	His	Ser	Leu	Phe	Asn	Ala
Leu	Ser	Ala	Lys	Thr	Lys	Leu	Gly	Glu	Asn	Gly	Arg	Phe	Phe	Leu	

<210>	216
<211>	2280
<212>	DNA
<213>	SHRIMP

<400> 216						
atgccaccta	aacataaaacc	taatactgct	ttaaaaaagc	atataattag	gaaccaacaa	60
cgaaaaaaag	aggatgacgc	cgaataccaga	ttccaacgaa	acatgggtca	ggaggtctct	120
aaattggcag	cgccaacttc	tagtaagaat	aggcaaaagg	gtaagattcg	tactagtataa	180
attctgtcaa	ggagtggcga	ttgtgttgct	ggcgaattgt	cgagatttaga	aaatgatgag	240
ggaaaaagag	acactgatca	agaaggtggt	ggaagaggag	gaggtaatga	agaagaagag	300
gaagggaaaag	aggaagggga	aggggaggaa	cagcagcggg	aagaaaaagga	agaacagagt	360
gaagagaaaag	aagaaaaaga	tggagaggag	gaagaagagg	agaattgtga	agacgaacat	420
gttacaccta	ctacatctgt	atcgaagcgg	gcaaaaacaa	tgaaaaagca	catatttctt	480
ccctctaaaa	aaagaaagag	aagtgcacaca	gaatcaaagg	cattagcagt	gcctgctgga	540
aaaatgatga	ccgtgtctcg	accattgaga	ggtgtctataa	cttcgggctc	tattttaggt	600
gtaaggagtg	aaaattcgcc	acaatacgat	tacgtgtcat	acttggcaga	cgaggcagtg	660
gtcaaggaga	aggctataca	gtacagaata	aggtctcttt	tagtaaacct	ttaaaagccc	720

```

aataaaacca aggccttttcc tacttcatca tctttattgt catctgaaca agggaagaaa 780
aagttcggag gaaaaagaac aaacactttt gttgtcacia atgtgggtgc tgaattggtg 840
aaagctcttc tggctaatag ttgttgggca atttcccata ggaaagatat acggtcagga 900
gagattcagt ggcaagaatt aagttcaaaa atattgaaat ctcttaacga tggaaatgca 960
acagagataa acaatttaat gagcagatc gtggaggatc gtatacagcg aactgtgaaa 1020
gaacgagtat actttgaaca acttgctact gtgtgtaata acctgtttgg tacacgtata 1080
ttaccgaaca agaactttga taagaatttt gtatctgtag cttctgataa ctctaacgcc 1140
actgtacgtg gcctatcaat acctcgttat tttagagcta taaataataa cgtgtgggta 1200
aaaatgtcat cgactatgga tcttttgggt ggaggaggga tgaggagtaa atcagagcat 1260
tctatctcca tgctggagaa atgtgcagct ggtgttttag caagagcttc agcaagaccg 1320
gttgaaaaaa tgataaagag tgcagtggag gaaacaagtc aagcatttaa tttatcgact 1380
ggagtgtttg tgcccaaaaca acaacaacag caacgtcagc aacaacaaca acaatttccc 1440
cctttccaac ctctccctt tctcttctt ccaccacagg cacttttcca ggttcaacaa 1500
cctacttacc aaggatactt aaatccttat tatcaatata accagtatta taaccatata 1560
gtcctcaac aattacaaca acagtaccg ctgtactttt taggcaacca atcacagccg 1620
ccgcctcagc tgcagcagca gcagcagcag cagcaaccac cacagccgcc caataatatt 1680
cctctctccc ctactctca acaacaatca ccagtaata ttctctctcc ccagcagcag 1740
cagcagcagc cctttccggt tcaactcatt tctagtcccc ctctctctcc tctataacct 1800
aatactgctc cttctccacc tatttcccg gttaagattg actctcgttc tactaccctt 1860
caacctccac ctacaccagt tctacccaag cctactctc ttctctctcc gtctacagca 1920
agagcagaag aagaaaacgc tactgatatg tcttttactg atatagactc tgagcttggc 1980
agtattgatt ttgatcttcc tcccgtact ccaggaggga acgttgaaga gataataaaa 2040
gcgcaacgtc aagctgtcaa ggaaacggga gtcagaggag aagaagaaga agaagaggag 2100
gcatttattg caccaattat tcgtcaaccg cgtacaccag gaaattttag agatgaactt 2160
ttagatgtca atgaatccat ctatggctca gacattgaac cagcagcagc agcagcagcg 2220
tttgactggg atatgggggtt agacgattta aatgggggatg aaccatatga atttgaataa 2280

```

<210> 217  
 <211> 757  
 <212> PRT  
 <213> SHRIMP

<400> 217

Met	Pro	Pro	Lys	His	Lys	Pro	Asn	Thr	Ala	Leu	Lys	Lys	His	Ile	Ile
1				5					10					15	
Arg	Asn	Gln	Gln	Arg	Lys	Lys	Glu	Asp	Ala	Glu	Ser	Arg		Phe	Gln
			20					25					30		
Arg	Asn	Met	Gly	Gln	Glu	Val	Ser	Lys	Leu	Asp	Ala	Pro	Thr	Ser	Ser
			35					40					45		
Lys	Asn	Arg	Gln	Arg	Arg	Lys	Ile	Arg	Thr	Ser	Lys	Ile	Leu	Ser	Arg
			50				55				60				
Ser	Gly	Asp	Cys	Val	Ala	Gly	Asp	Cys	Ser	Asp	Leu	Glu	Asn	Asp	Glu
					70					75					80
Gly	Lys	Arg	Asp	Thr	Asp	Gln	Glu	Gly	Gly	Gly	Arg	Gly	Gly	Gly	Asn
				85					90						95
Glu	Glu	Glu	Glu	Glu	Gly	Lys	Glu	Glu	Gly	Glu	Gly	Glu	Glu	Gln	Gln
			100					105					110		
Arg	Glu	Glu	Lys	Glu	Glu	Gln	Ser	Glu	Glu	Lys	Glu	Glu	Lys	Asp	Gly
			115					120					125		
Glu	Glu	Glu	Glu	Glu	Glu	Asn	Val	Glu	Asp	Glu	His	Val	Thr	Pro	Thr
			130			135					140				
Thr	Ser	Val	Ser	Lys	Arg	Ala	Lys	Gln	Met	Lys	Lys	His	Ile	Phe	Pro
					150					155					160
Pro	Ser	Lys	Lys	Arg	Lys	Arg	Ser	Asp	Thr	Glu	Ser	Lys	Ala	Val	Pro
				165					170					175	
Ala	Gly	Lys	Met	Met	Thr	Val	Ser	Arg	Pro	Leu	Arg	Gly	Ala	Ile	Thr
			180					185					190		
Ser	Gly	Ser	Ile	Leu	Gly	Val	Arg	Ser	Glu	Asn	Ala	Pro	Gln	Tyr	Asp
			195				200					205			
Tyr	Val	Ser	Tyr	Leu	Ala	Asp	Glu	Ala	Val	Val	Lys	Glu	Lys	Ala	Ile

	210					215					220					
Gln 225	Tyr	Arg	Ile	Arg	Ser 230	Leu	Leu	Ala	Asn	Leu 235	Leu	Lys	Ala	Asn	Lys 240	
Thr	Lys	Ala	Phe	Pro	Thr 245	Ser	Ser	Ser	Leu 250	Leu	Ser	Ser	Glu	Gln	Gly 255	
Lys	Lys	Lys	Phe	Gly	Gly 260	Lys	Arg	Thr 265	Asn	Thr	Phe	Val	Val	Thr	Asn	
Val	Gly	Ala	Glu	Leu	Val 275	Lys	Ala	Leu 280	Leu	Ala	Asn	Ser	Cys	Trp	Ala	
Ile	Ser	His	Arg	Lys	Asp 290	Ile	Arg	Ser 295	Gly	Glu	Ile	Gln	Trp	Gln	Glu	
Leu 305	Ser	Ser	Lys	Ile	Leu 310	Lys	Ser	Leu	Asn	Asp	Gly	Asn	Ala	Thr	Glu 320	
Ile	Asn	Asn	Leu	Met	Ser 325	Ser	Ile	Val	Glu	Asp	Arg	Ile	Gln	Arg	Thr 335	
Val	Lys	Glu	Arg	Val	Tyr 340	Phe	Glu	Gln	Leu	Ala	Thr	Val	Cys	Asn	Asn	
Leu	Phe	Gly	Thr	Arg	Ile 355	Leu	Pro	Asn	Lys	Asn	Phe	Asp	Lys	Asn	Phe	
Val	Ser	Val	Ala	Ser	Asp 370	Asn	Ser	Asn	Ala	Thr	Val	Arg	Gly	Leu	Ser	
Ile 385	Pro	Arg	Tyr	Phe	Arg 390	Ala	Ile	Asn	Asn	Asn	Val	Trp	Val	Lys	Met 400	
Ser	Ser	Thr	Met	Asp	Leu 405	Leu	Val	Gly	Gly	Gly	Met	Arg	Ser	Lys	Ser 415	
Glu	His	Ser	Ile	Ser	Met 420	Leu	Glu	Lys	Cys	Ala	Ala	Gly	Val	Leu	Ala	
Arg	Ala	Ser	Ala	Arg	Pro 435	Val	Glu	Lys	Met	Ile	Lys	Ser	Ala	Val	Glu	
Glu	Thr	Ser	Gln	Ala	Phe 450	Asn	Leu	Ser	Thr	Gly	Val	Phe	Val	Pro	Lys	
Gln 465	Gln	Gln	Gln	Gln	Arg 470	Gln	Gln	Gln	Gln	Gln	Gln	Phe	Pro	Pro	Phe 480	
Gln	Pro	Pro	Pro	Phe	Pro 485	Leu	Pro	Pro	Pro	Gln	Ala	Pro	Phe	Gln	Val 495	
Gln	Gln	Pro	Thr	Tyr	Gln 500	Gly	Tyr	Leu	Asn	Pro	Tyr	Tyr	Gln	Tyr	Asn	
Gln	Tyr	Tyr	Asn	Pro	Tyr 515	Ala	Pro	Gln	Gln	Leu	Gln	Gln	Gln	Tyr	Pro	
Leu	Tyr	Phe	Leu	Gly	Asn 530	Gln	Ser	Gln	Pro	Pro	Pro	Gln	Leu	Gln	Gln	
Gln 545	Gln	Gln	Gln	Gln	Gln 550	Pro	Pro	Gln	Pro	Pro	Asn	Asn	Ile	Pro	Pro 560	
Pro	Pro	Thr	Pro	Gln	Gln 565	Gln	Ser	Pro	Ser	Asn	Ile	Pro	Pro	Pro	Gln 575	
Gln	Gln	Gln	Gln	Gln	Pro 580	Phe	Pro	Val	Gln	Leu	Ile	Ser	Ser	Pro	Pro	
Pro	Pro	Pro	Pro	Ile	Pro 595	Asn	Thr	Ala	Pro	Ser	Pro	Pro	Ile	Ser	Arg	
Val	Arg	Phe	Asp	Ser	Arg 610	Ser	Thr	Thr	Pro	Gln	Pro	Pro	Pro	Thr	Pro	
Val 625	Leu	Pro	Lys	Pro	Thr 630	Pro	Leu	Pro	Pro	Pro	Ser	Thr	Ala	Arg	Ala 640	
Glu	Glu	Glu	Asn	Ala	Thr 645	Asp	Met	Ser	Phe	Thr	Asp	Ile	Asp	Ser	Glu 655	
Leu	Gly	Ser	Ile	Asp	Phe 660	Asp	Leu	Pro	Pro	Ala	Thr	Pro	Gly	Arg	Asn	
Val	Glu	Glu	Ile	Ile	Lys 675	Ala	Gln	Arg	Gln	Ala	Val	Lys	Glu	Thr	Gly	
Val	Arg	Gly	Glu	Glu	Glu 690	Glu	Glu	Glu	Glu	Ala	Phe	Ile	Ala	Pro	Ile 700	

WO 01/38351

307

PCT/US00/28888

Ile Arg Gln Pro Arg Thr Pro Gly Asn Phe Arg Asp Glu Leu Leu Asp  
 705 710 715 720  
 Val Asn Glu Ser Ile Tyr Gly Ser Asp Ile Glu Pro Ala Ala Ala Ala  
 725 730 735  
 Ala Ala Phe Asp Trp Asp Met Gly Leu Asp Asp Leu Asn Gly Asp Glu  
 740 745 750  
 Pro Tyr Glu Phe Glu  
 755

<210> 218  
 <211> 3177  
 <212> DNA  
 <213> SHRIMP

<400> 218  
 atggaagggtg taacatccat agttgcagct gctgtcccag aagttgcaat cttgatcacc 60  
 gacctaattgg gaggcagaaa taataagagg tccacctatg aacgaattgt gggatatagt 120  
 ggagaaagtgt gtgatcttct ggaggccata ctggacatct gcaaccgcaa ttcttatagg 180  
 gatgaacttt tggagggaga aacagttgtg attaatccaa ccggtctttt gaaagagata 240  
 tctctatttaa tgaagaaagc cctcgacatg aacatcaaga tgaaaagtaa cgacgacca 300  
 gttccattta caactcttga ccaaaacgaa caagaattca ttggacattt aaagagttgc 360  
 aaaaaacaag atgggcccagc atacaaggat ctaattcatc gtatttatag tggcatgttt 420  
 gtcatgaaaa atacgcgctt gatgttggtt gaaattatac gtggaaatgc aggcgatgca 480  
 gtggaggaga agaatgcgtt gtgtgaggcg tacgccgaga tgatatcaga tatggatttg 540  
 atacgcattt tcttactcct tgttgccatc aaacgagatc agaacaagaa gcaccgccac 600  
 atgaaaagtgt tcatattatga agatgtggta gtttctctca acactttgaa ggatgtgttt 660  
 cataaggagt ggtacatgtg gcccttctct gctttgcaag tggggacaaa aataagagac 720  
 gctagaacgt ttagtggttt atttggttct gatatgcag aaggaagaaa caatgacagg 780  
 atttgggaaa atatggcctt ctccgtaaca gaagcatttt taagtgtgct atccaccaac 840  
 aaccatcata ataagggtca tcttcgcag tacgctgcca ggcccgtgta tgacgctatg 900  
 gaattgtgac cccaagaatt acaccatatt ctttttgga caaaaatagc caagatgatt 960  
 gatatagtgt acagatactc tatctacaat gttccctatc ttttggcagc agatactgaa 1020  
 agagtagagg agccaaagaa atccgtaatg tcgccttctg gtctaataat ctccccgaac 1080  
 gcttcgctct tagaaaatac accactgtcc ctgggtatct gacacggcat accttcagcc 1140  
 agaaaactag gatctttcat cctcgaacat gaaaacgctg aaaaatagca tctagaagaa 1200  
 gcaatcgcca aatgcattgt tagccagaca ctgcaagaag agtcttgagg agaaagtcaa 1260  
 gcggcgatgg tgtaccaacc atctgacgaa gtggaggtaa tacaagcaca cgttacaaag 1320  
 atattgagtg ggaataccac caataaaaca tgtgggtgtg gctacgcgga tttggacatg 1380  
 aaacccaaat tcttcaactg ttctcatgaa aacatgaaag cctcttacga ttatttcccc 1440  
 gtacatgctt tcatggacac ttttgaagca cgacaagaaa cgtgttctgc aaaactatgc 1500  
 cctgattgta ctataaaaca cctaattgat gtgtacgaaa aggtctcggc tggtagtgaa 1560  
 aaactaaagg acgttttttag gtgcccattg ttgtggaga atatatgtaca gtttatagg 1620  
 agatgtcacg aattttcatc tttattcgag agggcgattt tagcaggaga aaatgtcgac 1680  
 ccagaatata ttgcagccaa caaacttctc atcaccgaac tcatcaagcg agcagagaaa 1740  
 tgcttctata ctgtggaact tcttcaggcc gaattcatgg agatgtgtaa gatggacaag 1800  
 gactttgcac tagacaagga ttcaaaattt actgtgtgga ataatagatt cagaccgccc 1860  
 gttaaactct ttaagatggt ggaggagaa actggtgata gcaaatgttc cctgatctgt 1920  
 acccagtggt tactacccaa tgtatgtgac cagcctaag agatggaaga cattgttaaa 1980  
 gtggatgtac ctccctcag attgcccgtac cctccctccg aacaactgga agactattac 2040  
 ttccaggatg tagaagatgc tgaatttgat gaccctccca ccgatgagct agtaagatat 2100  
 gatactgggt ccggaactcca taaatggccc atgagattgt cctgcggttt cttggcgatc 2160  
 aattttgtac ctccaaatga agaggtgaca aattgtaggc aggtctgttt tatactaaaa 2220  
 agaactcccg aaaagaagat tcgaggatgg aaccagagt ctccggaggg caaggttctt 2280  
 ctacgcctag ccaactggca ctcgactgat cgtatgccag agaatatgaa ggggcttctc 2340  
 aacgacatat ctgtaatcca caacaccagg gagaggtttc agaatagggt taaagtccat 2400  
 tacttaaaac gtgttttttg gggttttgat gatagagatt ttgaacaagt ttaggtgtg 2460  
 agtataccct taattgcaac ctatttttat gtctatgaaa aattgaacca tgagagtgca 2520  
 ttaggattat gggcgaaaaat gtttgtcaag aatcttatag gggaaatggt gttggagcgg 2580  
 ccagaatgtg tgttccatag agcccattct tttgtccttc attgtgtaga taggcgtgcc 2640  
 ttatcgggta tcagacctaa ccaaggtgca aaaatggaaa ttgttaaaac ggtaaatatt 2700

gtacgtcaaa acatgacatc tgaatcaata aaggatccag tgttcactgt ggatgaaaag 2760  
aggacattag aatggaaggt tgagaaagaa ggacaggaaa ttaagactgt aaaatgccct 2820  
aaatgtaaga cacctaacat aaaattaggt ggatgtatta ccatgacttg ttatgattgt 2880  
agcggtagaa gagatgggta ccctacagtt ttttgttgga tatgtgaaga tgaataaact 2940  
aaccctgacc atatacta atactataat tgaccataaa ctattgtatt ctgattgtaa gtctacaaag 3000  
gctgcttttag aaaaggtgta taattgtact ctatgttgtc tagctttaag aaaatgtagt 3060  
gatagttatt tgtccaaaca aagaggagga ggagaagaag aagagattga gatatatgta 3120  
atggaagatg gttttgaatt tgatgtacat actaaaacag cgggtgcctac aaaataa 3177

<210> 219

<211> 1050

<212> PRT

<213> SHRIMP

<400> 219

Met	Glu	Gly	Val	Thr	Ser	Ile	Val	Ala	Ala	Ala	Val	Pro	Glu	Val	Ala
1				5				10						15	
Ile	Leu	Ile	Thr	Asp	Leu	Met	Gly	Gly	Arg	Asn	Asn	Lys	Arg	Ser	Thr
			20				25					30			
Tyr	Glu	Arg	Ile	Val	Gly	Ile	Val	Gly	Glu	Ser	Gly	Asp	Leu	Leu	Glu
		35					40					45			
Ala	Ile	Leu	Asp	Ile	Cys	Asn	Arg	Asn	Ser	Tyr	Arg	Asp	Glu	Leu	Leu
	50				55						60				
Glu	Gly	Glu	Thr	Val	Val	Ile	Asn	Pro	Thr	Gly	Leu	Leu	Lys	Glu	Ile
	65				70					75				80	
Ser	Leu	Leu	Met	Lys	Lys	Ala	Leu	Asp	Met	Asn	Ile	Lys	Met	Lys	Ser
			85					90						95	
Asn	Asp	Asp	Pro	Val	Pro	Phe	Thr	Thr	Leu	Asp	Gln	Asn	Glu	Gln	Glu
			100					105						110	
Phe	Ile	Gly	His	Leu	Lys	Ser	Cys	Lys	Lys	Gln	Asp	Gly	Pro	Ala	Tyr
		115					120					125			
Lys	Asp	Leu	Ile	His	Arg	Ile	Tyr	Ser	Gly	Met	Phe	Val	Met	Lys	Asn
	130					135					140				
Thr	Arg	Leu	Met	Leu	Asp	Glu	Ile	Ile	Arg	Gly	Asn	Ala	Gly	Asp	Ala
	145				150					155				160	
Val	Glu	Glu	Lys	Asn	Ala	Leu	Cys	Glu	Ala	Tyr	Ala	Glu	Met	Ile	Ser
			165					170						175	
Asp	Met	Asp	Leu	Ile	Arg	Ile	Phe	Leu	Leu	Leu	Val	Ala	Ile	Lys	Arg
			180					185						190	
Asp	Gln	Asn	Lys	Lys	His	Arg	His	Met	Lys	Ser	Val	Ile	Tyr	Glu	Asp
	195						200					205			
Val	Val	Val	Ser	Leu	Asn	Thr	Leu	Lys	Asp	Val	Phe	His	Lys	Glu	Trp
	210					215					220				
Tyr	Met	Trp	Pro	Phe	Ser	Ala	Leu	Gln	Val	Gly	Thr	Lys	Ile	Arg	Asp
	225				230					235				240	
Ala	Arg	Thr	Phe	Ser	Val	Leu	Phe	Gly	Ser	Asp	Met	His	Glu	Gly	Arg
			245					250						255	
Asn	Asn	Asp	Arg	Ile	Trp	Glu	Asn	Met	Ala	Phe	Ser	Val	Thr	Glu	Ala
			260					265						270	
Phe	Leu	Ser	Gly	Pro	Ser	Thr	Asn	Asn	His	Tyr	Asn	Lys	Gly	His	Leu
		275					280					285			
Arg	Met	Tyr	Ala	Ala	Arg	Pro	Val	Tyr	Asp	Ala	Met	Glu	Tyr	Val	Pro
	290					295						300			
Gln	Glu	Leu	His	His	Ile	Leu	Phe	Gly	Thr	Lys	Ile	Ala	Lys	Met	Ile
	305				310					315				320	
Asp	Ile	Val	Tyr	Arg	Tyr	Ser	Ile	Tyr	Asn	Val	Pro	Tyr	Leu	Leu	Ala
			325					330						335	
Ala	Asp	Thr	Glu	Arg	Val	Glu	Glu	Pro	Lys	Lys	Ser	Val	Met	Ser	Pro
			340					345					350		
Ser	Gly	Leu	Ile	Ile	Ser	Pro	Asn	Ala	Ser	Leu	Leu	Glu	Asn	Thr	Pro
	355						360					365			



WO 01/38351

309

PCT/US00/28888

Leu	Ser	Leu	Val	Ser	Arg	His	Gly	Ile	Pro	Ser	Ala	Arg	Lys	Leu	Gly
370						375					380				
Ser	Phe	Ile	His	Glu	Asn	Ala	Glu	Asn	Met	His	Leu	Glu	Glu	Ala	Ile
385					390					395					400
Ala	Lys	Cys	Met	Val	Ser	Gln	Thr	Leu	Gln	Glu	Glu	Ser	Trp	Gly	Glu
			405						410					415	
Ser	Gln	Ala	Ala	Met	Val	Tyr	Gln	Pro	Ser	Asp	Glu	Val	Glu	Val	Ile
			420					425					430		
Gln	Ala	His	Val	Thr	Lys	Ile	Leu	Ser	Gly	Asn	Thr	Thr	Asn	Lys	Thr
		435					440					445			
Cys	Gly	Leu	Cys	Tyr	Ala	Asp	Leu	Asp	Met	Lys	Pro	Lys	Phe	Phe	Asn
450						455					460				
Cys	Ser	His	Glu	Asn	Met	Lys	Ala	Ser	Tyr	Asp	Tyr	Phe	Pro	Val	His
465					470					475					480
Ala	Phe	Met	Asp	Thr	Phe	Glu	Ala	Arg	Gln	Glu	Thr	Cys	Ser	Ala	Lys
				485					490						495
Leu	Cys	Pro	Asp	Cys	Thr	Ile	Lys	His	Leu	Met	Tyr	Val	Tyr	Glu	Lys
			500					505					510		
Val	Ser	Ala	Gly	Ser	Glu	Lys	Leu	Lys	Asp	Val	Phe	Arg	Cys	Pro	Cys
		515					520					525			
Cys	Gly	Glu	Tyr	Met	Val	Gln	Phe	Ile	Gly	Arg	Cys	His	Glu	Phe	Ser
530						535					540				
Ser	Leu	Phe	Glu	Arg	Ala	Ile	Leu	Ala	Gly	Glu	Asn	Val	Asp	Pro	Glu
545					550					555					560
Tyr	Ile	Ala	Ala	Asn	Lys	Leu	Leu	Ile	Thr	Glu	Leu	Ile	Lys	Arg	Ala
				565					570						575
Glu	Lys	Cys	Phe	Tyr	Thr	Val	Glu	Leu	Gln	Ala	Glu	Phe	Met	Glu	
			580					585				590			
Met	Cys	Lys	Met	Asp	Lys	Asp	Phe	Ala	Leu	Asp	Lys	Asp	Ser	Lys	Phe
		595					600					605			
Thr	Val	Val	Asp	Asn	Arg	Phe	Arg	Pro	Pro	Val	Lys	Leu	Phe	Lys	Met
610						615					620				
Val	Glu	Gly	Glu	Thr	Gly	Asp	Ser	Lys	Cys	Ser	Leu	Ile	Cys	Thr	Gln
625					630					635					640
Cys	Leu	Leu	Pro	Asn	Val	Cys	Asp	Gln	Pro	Asn	Glu	Met	Glu	Asp	Ile
				645					650					655	
Val	Thr	Val	Asp	Val	Pro	Pro	Pro	Val	Leu	Pro	Tyr	Pro	Pro	Pro	Glu
			660					665					670		
Gln	Leu	Glu	Asp	Tyr	Tyr	Phe	Gln	Asp	Val	Glu	Asp	Ala	Glu	Phe	Asp
		675					680					685			
Asp	Pro	Pro	Thr	Asp	Glu	Leu	Val	Arg	Tyr	Asp	Thr	Gly	Pro	Gly	Leu
690						695					700				
His	Lys	Trp	Pro	Met	Arg	Leu	Ser	Cys	Gly	Phe	Leu	Asn	Phe	Val	Pro
705					710					715					720
Pro	Asn	Glu	Glu	Val	Thr	Asn	Cys	Arg	Gln	Ala	Val	Ser	Ile	Leu	Lys
				725					730						735
Arg	Thr	Pro	Glu	Lys	Lys	Ile	Arg	Gly	Trp	Asn	Pro	Glu	Ser	Pro	Glu
			740					745					750		
Gly	Lys	Val	Leu	Leu	Ala	Asn	Trp	His	Ser	Thr	Asp	Arg	Met	Pro	Glu
		755					760					765			
Asn	Met	Lys	Gly	Leu	Leu	Asn	Asp	Ile	Ser	Val	Ile	His	Asn	Thr	Arg
770						775					780				
Glu	Arg	Phe	Gln	Asn	Arg	Val	Lys	Val	His	Tyr	Leu	Asn	Ser	Val	Phe
785					790					795					800
Gly	Gly	Phe	Asp	Asp	Arg	Asp	Phe	Glu	Gln	Val	Val	Gly	Val	Ser	Ile
				805					810						815
Pro	Leu	Ile	Ala	Thr	Tyr	Phe	Tyr	Val	Tyr	Glu	Lys	Leu	Asn	His	Glu
				820				825					830		
Ser	Ala	Leu	Gly	Leu	Trp	Ala	Lys	Met	Phe	Val	Lys	Asn	Leu	Ile	Gly
		835					840					845			
Glu	Met	Val	Leu	Glu	Arg	Pro	Glu	Cys	Val	Phe	His	Arg	Ala	His	Ser

WO 01/38351

310

PCT/US00/28888

```

      850              855              860
Phe Val Leu His Cys Val Asp Arg Arg Ala Leu Ser Gly Ile Arg Pro
865              870              875              880
Asn Gln Gly Ala Lys Met Glu Ile Val Lys Gln Val Asn Ile Val Arg
      885              890              895
Gln Asn Met Thr Ser Glu Ser Ile Lys Asp Pro Val Phe Thr Val Asp
      900              905              910
Glu Lys Arg Thr Leu Glu Trp Lys Val Glu Lys Glu Gly Gln Glu Ile
      915              920              925
Lys Thr Val Lys Cys Pro Lys Cys Lys Thr Pro Asn Ile Lys Leu Gly
      930              935              940
Gly Cys Ile Thr Met Thr Cys Tyr Asp Cys Ser Gly Arg Arg Asp Gly
945              950              955              960
Tyr Pro Thr Val Phe Cys Trp Ile Cys Glu Asp Glu Ile Thr Asn Pro
      965              970              975
Asp His Ile Leu Ile Asp His Lys Leu Lys Asp Cys Lys Ser Thr Lys
      980              985              990
Ala Ala Leu Glu Lys Val Tyr Asn Cys Thr Leu Cys Cys Leu Ala Leu
      995              1000              1005
Arg Lys Cys Ser Asp Ser Tyr Leu Ser Lys Gln Arg Gly Gly Gly Glu
      1010              1015              1020
Glu Glu Glu Ile Glu Ile Tyr Val Met Glu Asp Gly Phe Glu Phe Asp
1025              1030              1035              1040
Val His Thr Lys Thr Ala Val Pro Thr Lys
      1045              1050

```

<210> 220  
 <211> 2547  
 <212> DNA  
 <213> SHRIMP

```

<400> 220
atgggttcta accagcaaca atcattcatc tcaaagagga atggcactaa gcaagaaatt 60
agtcttgaaa agataatcaa gaggattgaa aatgcctgtc ttccagtcaa ccagtatgtg 120
cccaagcttg acaagaacgc aattaaccct caagaacttg catctcacat catggaccgt 180
cttcccgcta ccatctcctt ccaagaaatg gacgattttc tggccgatta tgcaaagaca 240
aaaattgttg accaccctga ttttgaaaaa ctggcaggaa gattcatctg ttcgaacatc 300
cacaaaaaca ccaaagagtg gaatagtttt agtgcaacaa ctcaagaaat gaggcacgca 360
atccaccctg gaactggtaa accagcatca gtagttaatg atacctacta tgaaaatgtt 420
atggccaatg ctgaaattct cgatgctgtt atcgattaca aaatggatta tctcttcacc 480
tgtttcggac tgaggacgct agaatactct tatttgatca agattggttc cccactgat 540
agaaaagaaga gaatcttggt tgagcgccct caggacatga ttatgcgtgt ggctgtcggc 600
attcacggat cagacatcaa atctgtcatt gaaacatatg atctcatgtc gaggcactat 660
tttaccctatg ctccccccac actgtttaat tgtggaacag tcacacccca actttcctcg 720
tgcttccttc tgggccttca agatgatagc attgagggtt tttatgatac tcttaaggag 780
gcggcaatca tctccaagac tgctggagga ctccggcatc actttcatga tttgagggca 840
aaagggaagcc ccatttcgtc atggagtggg acccaccctg gtctcatggc gttcctccaa 900
atctttaacg tctctgtgaa aaagggttagc cagggaggag acaagaggag aggagctgca 960
gccatctata tttctgattg gcatctggac gtgaaggact ttattgactg cagaaagaat 1020
gccggtaatg aagatttgag gacgagggat cttttcccag ctatctgggt atctgatctc 1080
ttcatggaga gagtgaaggc tgggaagaat tgggtccctga tgtgccccca cgagtgcctt 1140
ggcctttccg acgtccatgg cgaagagttt aaggcgttgt acgagaaata tgaggctgaa 1200
ggcaaaggta aagagggtgt gaaggcacgt gcattattcg accaaattaa ttctgcacgt 1260
atcgaaactg gaacacctta tgtgtgcttt aaggatacca tcaatagaaa gtctaaccaa 1320
gaaaatgtcg gcatcatcaa gtcttcaaat ttgtgcactg aaattgtcca gtacagtgat 1380
tcggaggaaa ctgcagtgtg caatttggct tctatcgag tcaacaagtt tgtgaagtat 1440
tctcccatcc cttccctaag gcctatgtt gattaccggg agatgaagag ggttgtaaaa 1500
atcatgacca gaaatctcga caaggtgatt gatgtcaatt tctatgcagt tgacaagacc 1560
cgcatttcca atatgaaaac taggccaatg ggattgggtg tgcagggact agcagatttg 1620
ttcttcaaac tcagaatccc cttcgaatct gaagaagcgg cactaattaa caaggaggatt 1680

```

t ttgaaacta	tatactatgg	tgcttggaa	gcttcattgt	aaattgccaa	agaaaagggg	1740
gaaacatatg	agctgtttga	gggtagtcct	ttgagcaaag	ggattttcca	atttgacatg	1800
gggaagggaa	atattaagaa	tagggacata	tatttcaact	ctttgccaat	tcacgattgg	1860
gagcaattga	gaagggacat	tatgaagtat	gggtgtccaa	aatcaattgt	tgttgctccc	1920
atgcctactg	ctcacactgc	acagatcctg	ggcaactctg	atccctttga	gcctttaacg	1980
tccaacatgt	ataatcgtaa	tgtactttca	ggatcattcc	aagtggtgaa	cgaatatgtg	2040
attagagagc	tcataaaact	gggagaatgg	aattcagtaa	ctaacacagag	gattatggca	2100
agtgggtgat	ctattcgacg	gcttcctaatt	atccctaatt	cgaccaagga	actattcaaa	2160
actgttatgg	aaattaatcc	tcgtactact	ttggacattg	ctattcgagag	aggtagtttt	2220
ctgtgaccaag	ctcaatccct	caacttgttt	gtggaagaac	cgaactcag	caaggtgcgg	2280
tcaatgacta	tgtatgcattg	ggagaagggg	atcaagactc	tctattatct	acgcacaaag	2340
ggcgagctta	gagctgtcca	gttcacagtc	gacaagaatg	tactccaaga	agtcaagaag	2400
gaagctccct	ctctctgttg	agctttttct	gctcctgtcc	gagaagaaga	agaggagaag	2460
aagtcctcta	ttgttgttcc	agatcctgct	gctgctcttt	tatgttctat	caataaccct	2520
ggtgcttctg	aaatgtatcc	ttcctaag				2547

```
<210> 221
<211> 842
<212> PRT
<213> SHRIMP
```

<400> 221

Met 1	Gly	Ser	Asn	Gln	Gln	Gln	Ser	Phe	Ile	Ser	Lys	Arg	Asn	Gly	Thr
Lys	Gln	Glu	Ile	Ser	Leu	Glu	Lys	Ile	Ile	Lys	Arg	Ile	Glu	Asn	Ala
Cys	Leu	Pro	Val	Asn	Gln	Tyr	Val	Pro	Lys	Leu	Asp	Lys	Asn	Ala	Ile
Asn	Pro	Gln	Glu	Leu	Ala	Ser	His	Ile	Met	Asp	Arg	Leu	Pro	Ala	Thr
Ile	Ser	Phe	Gln	Glu	Met	Asp	Asp	Phe	Leu	Ala	Asp	Tyr	Ala	Lys	Thr
Lys	Ile	Val	Asp	His	Pro	Asp	Phe	Gly	Lys	Leu	Ala	Gly	Arg	Phe	Ile
Cys	Ser	Asn	Ile	His	Lys	Asn	Thr	Lys	Glu	Trp	Asn	Ser	Phe	Ser	Ala
Thr	Thr	Gln	Lys	Leu	Arg	His	Ala	Ile	His	Pro	Gly	Thr	Gly	Lys	Pro
Ala	Ser	Val	Val	Asn	Asp	Thr	Tyr	Tyr	Glu	Asn	Val	Met	Ala	Asn	Ala
Glu	Ile	Leu	Asp	Ala	Val	Ile	Asp	Tyr	Lys	Met	Asp	Tyr	Leu	Phe	Thr
Cys	Phe	Gly	Leu	Arg	Thr	Leu	Glu	Tyr	Ser	Tyr	Leu	Ile	Lys	Ile	Gly
Ser	Pro	Thr	Asp	Arg	Lys	Lys	Arg	Ile	Leu	Val	Glu	Arg	Pro	Gln	Asp
Met	Ile	Met	Arg	Val	Ala	Val	Gly	Ile	His	Gly	Ser	Asp	Ile	Lys	Ser
Val	Ile	Glu	Thr	Tyr	Asp	Leu	Met	Ser	Arg	His	Tyr	Phe	Thr	His	Asp
Thr	Leu	Phe	Asn	Cys	Gly	Thr	Val	Thr	Pro	Gln	Leu	Ser	Ser	Cys	Phe
Leu	Leu	Gly	Leu	Gln	Asp	Asp	Ser	Ile	Glu	Gly	Ile	Tyr	Asp	Thr	Leu
Lys	Glu	Ala	Ala	Ile	Ile	Ser	Lys	Thr	Ala	Gly	Gly	Leu	Gly	Ile	His
Phe	His	Asp	Leu	Arg	Ala	Lys	Gly	Ser	Pro	Ile	Ser	Ser	Trp	Ser	Gly
Thr	His	Pro	Gly	Leu	Met	Ala	Phe	Leu	Gln	Ile	Phe	Asn	Val	Ser	Val
Lys	Lys	Val	Ser	Gln	Gly	Gly	Asp	Lys	Arg	Arg	Gly	Ala	Ala	Ala	Ile

305					310					315				320	
Tyr	Ile	Ser	Asp	Trp	His	Leu	Asp	Val	Lys	Asp	Phe	Ile	Asp	Cys	Arg
				325					330					335	
Lys	Asn	Ala	Gly	Asn	Glu	Asp	Leu	Arg	Thr	Arg	Asp	Leu	Phe	Pro	Ala
			340					345					350		
Ile	Trp	Val	Ser	Asp	Leu	Phe	Met	Glu	Arg	Val	Lys	Ala	Gly	Lys	Asn
		355					360					365			
Trp	Ser	Leu	Met	Cys	Phe	Cys	Pro	Gly	Leu	Ser	Asp	Val	His	Gly	Glu
	370					375					380				
Glu	Phe	Lys	Ala	Leu	Tyr	Glu	Lys	Tyr	Glu	Ala	Glu	Gly	Lys	Gly	Lys
385					390				395						400
Glu	Val	Val	Lys	Ala	Arg	Ala	Leu	Phe	Asp	Gln	Ile	Asn	Ser	Ala	Arg
				405					410					415	
Ile	Glu	Thr	Gly	Thr	Pro	Tyr	Val	Cys	Phe	Lys	Asp	Thr	Ile	Asn	Arg
			420					425					430		
Lys	Ser	Asn	Gln	Glu	Asn	Val	Gly	Ile	Ile	Lys	Ser	Ser	Asn	Leu	Cys
		435					440					445			
Thr	Glu	Ile	Val	Gln	Tyr	Ser	Asp	Ser	Glu	Glu	Thr	Ala	Val	Cys	Asn
	450					455					460				
Leu	Ala	Ser	Ile	Ala	Val	Asn	Lys	Phe	Val	Lys	Tyr	Ser	Pro	Ile	Pro
465					470				475						480
Ser	Leu	Arg	Pro	Tyr	Val	Asp	Tyr	Arg	Glu	Met	Lys	Arg	Val	Val	Lys
				485					490					495	
Ile	Met	Thr	Arg	Asn	Leu	Asp	Lys	Val	Ile	Asp	Val	Asn	Phe	Tyr	Ala
			500					505					510		
Val	Asp	Lys	Thr	Arg	Ile	Ser	Asn	Met	Lys	Trp	Met	Gly	Leu	Gly	Val
	515						520					525			
Gln	Gly	Leu	Ala	Asp	Leu	Phe	Phe	Lys	Leu	Arg	Ile	Pro	Phe	Glu	Ser
	530					535					540				
Glu	Glu	Ala	Ala	Leu	Ile	Asn	Lys	Arg	Ile	Phe	Glu	Thr	Ile	Tyr	Tyr
545					550					555					560
Gly	Ala	Leu	Glu	Ala	Ser	Cys	Glu	Ile	Ala	Lys	Glu	Lys	Gly	Glu	Thr
				565					570					575	
Tyr	Glu	Leu	Phe	Glu	Gly	Ser	Pro	Leu	Ser	Lys	Gly	Ile	Phe	Gln	Phe
			580					585					590		
Asp	Met	Gly	Lys	Glu	Asn	Ile	Lys	Asn	Arg	Asp	Ile	Tyr	Phe	Asn	Ser
		595					600					605			
Leu	Pro	Ile	His	Asp	Trp	Glu	Gln	Leu	Arg	Arg	Asp	Ile	Met	Lys	Tyr
	610					615					620				
Gly	Val	His	Asn	Ser	Met	Phe	Val	Ala	Pro	Met	Pro	Thr	Ala	Ser	Thr
625					630					635					640
Ala	Gln	Ile	Leu	Gly	Asn	Ser	Glu	Ser	Phe	Glu	Pro	Leu	Thr	Ser	Asn
				645					650					655	
Met	Tyr	Asn	Arg	Asn	Val	Leu	Ser	Gly	Ser	Phe	Gln	Val	Val	Asn	Glu
			660					665					670		
Tyr	Val	Ile	Arg	Glu	Leu	Ile	Lys	Leu	Gly	Glu	Trp	Asn	Ser	Val	Thr
		675					680					685			
Lys	Gln	Arg	Ile	Met	Ala	Ser	Gly	Gly	Ser	Ile	Gln	Thr	Leu	Pro	Asn
	690					695					700				
Ile	Pro	Lys	Ser	Thr	Lys	Glu	Leu	Phe	Lys	Thr	Val	Trp	Glu	Ile	Asn
705					710					715					720
Pro	Arg	Thr	Thr	Leu	Asp	Met	Ala	Ile	Gln	Arg	Gly	Met	Phe	Val	Asp
				725					730					735	
Gln	Ala	Gln	Ser	Leu	Asn	Leu	Phe	Val	Glu	Glu	Pro	Glu	Leu	Ser	Lys
				740				745					750		
Val	Arg	Ser	Met	Thr	Met	Tyr	Ala	Trp	Glu	Lys	Gly	Ile	Lys	Thr	Leu
		755					760					765			
Tyr	Tyr	Leu	Arg	Thr	Lys	Gly	Ala	Ala	Arg	Ala	Val	Gln	Phe	Thr	Val
	770					775					780				
Asp	Lys	Asn	Val	Leu	Gln	Glu	Val	Lys	Lys	Glu	Ala	Pro	Ser	Pro	Val
785					790					795					800

WO 01/38351

313

PCT/US00/28888

Ala	Ala	Phe	Ser	Ala	Pro	Val	Arg	Glu	Glu	Glu	Glu	Glu	Lys	Lys	Ser
			805					810					815		
Ser	Ile	Val	Val	Pro	Asp	Pro	Ala	Ala	Ala	Leu	Leu	Cys	Ser	Ile	Asn
		820						825					830		
Asn	Pro	Gly	Ala	Cys	Glu	Met	Cys	Ser	Ser						
		835					840								

<210> 222  
 <211> 222  
 <212> DNA  
 <213> SHRIMP

<400> 222  
 atgggtcatatc gactctgctt cttggaaaagt atcacgtgct ttgtatatgg catcatggca 60  
 ccactatctt tggacacaaa taccgattat ttatcccaca aaaaggatac taataaaaaa 120  
 atacagatgc agataaaactt tattccttat tcaaatatgc acgtatatat agcagggtgtg 180  
 tacacatttc atgaaaaaaa ggggttaaca tatcaacaat at 222

<210> 223  
 <211> 74  
 <212> PRT  
 <213> SHRIMP

Met	Val	Ile	Arg	Leu	Cys	Phe	Leu	Glu	Ser	Ile	Thr	Cys	Phe	Val	Tyr
1				5					10					15	
Gly	Ile	Met	Ala	Pro	Leu	Ser	Leu	Asp	Thr	Asn	Thr	Asp	Tyr	Leu	Ser
			20					25					30		
His	Lys	Lys	Asp	Thr	Asn	Lys	Lys	Ile	Gln	Met	Gln	Ile	Asn	Phe	Ile
		35				40					45				
Pro	Tyr	Ser	Asn	Met	His	Val	Tyr	Ile	Ala	Gly	Val	Tyr	Thr	Phe	His
	50					55					60				
Glu	Lys	Lys	Gly	Leu	Thr	Tyr	Gln	Gln	Tyr						
65						70									

<210> 224  
 <211> 1848  
 <212> DNA  
 <213> SHRIMP

<400> 224  
 atggaatata ttggagagaa aaacaataac cctgttagta atgaaagtgt atcagaaaaag 60  
 gagttaaaac taagatcgct attcctgatg atcggaaga aaacaagtaa atatgagcaa 120  
 gtcattgggtg tttatgaagc tatcgaatct ataagacaaa gcgaattgtc cgaagacaca 180  
 tttgttgtac atgtgaagaa agataaacia ctcaaatctg caagagggtt aaaaagatta 240  
 caagaattgg tagaagatga ctctttaaga attgaacgga taagttgtgc ccctcctgaa 300  
 cctggacatt tattcaaaga tgatgctggg cacgttactg acgaggaatg gcttgcaacg 360  
 caagaagaag acgtgcgtaa aatcaatata atagtcaagg aaaaattaaa acgaaaagac 420  
 aaggacttta aattcagtc attatacagg tacatgagca atagtctttc tgaagcagta 480  
 gaaaaaaaac acgattgtat gataataagt tccgatttct taatcgggtt aggtttcagt 540  
 acaatgaacg tcacgcacgc tttaaagtca atggagagaa ctatgcagaa acatggtttc 600  
 aaggatatga tggccccatt gggttgaaatt tgtcaccgta cccattacaa aggagaatat 660  
 atagccaatc ctattttcaa gagccattct tcacattgct tgattgttcc cttgtttatg 720  
 gtggcgggag tttttgcaag gagcgcacac cttctgctg caagcattga aatgtacctt 780  
 tcgacactag cctacgctgt tatcttatac agtgatgaaa agcaacgcca gatacgcgaa 840  
 gagttggcta ggaaaaattt acaaaataaa gaggaactag aaaaccaggc cgaaaagacc 900  
 acaaaagtgtg aaaaaggact agaaacacaa gtagttaaga ccacaaaagt tgaaaaggaa 960  
 ctagaaacac aagtagtaaa gaaagaggag taaaaaaact cgtatatcga aactgaacaa 1020  
 cttttcaaag tctctgagga aaaaaaagaa tctctcagaa atgtacacaa gaaatcttcc 1080

aatgcgacct	tcagatacga	cagcggctct	tgtctcgtct	tctcgatata	ctctacagaa	1140
ttctacttgt	tgtcaggac	agacaaaagt	ggttcctttg	aaactgcaac	agaaaaatggt	1200
ttgaggtata	ttttctcccc	cataaacaag	aaaagggata	ctgcaggat	gagggcccga	1260
ctcatcatgg	ccgtgactgg	ctgtgacgca	ccatctgcctt	gcaacgacag	tatcaaacac	1320
caaaacaagt	tcaaggtatt	aaaatgtaac	cgatcaagta	tagttttcca	gaccctcca	1380
agtgatgaag	atttgaagg	tattgtacaa	aaagtgcag	gtctcgatat	ccgaattctt	1440
atgaatgatg	gcaccgtcta	tcaagatggg	cacagggatg	acatctcttc	gcctcaagaa	1500
cttgatgaag	aaaatatgac	ccaatttgaa	attgaacaac	aaaggaaagt	ccattccatg	1560
atggagaaac	catcaaaaat	tgtcactgag	tacaacaagg	aaagacattt	gaccacaagg	1620
gaagctcgta	cgagaacaaa	gaccgaaaag	tggtttgaga	aggtaaagaa	gagggaggaa	1680
caaaagaaag	gagaaaaatg	agaacagtct	accagtgaa	aggagcaaga	gggagtaaaa	1740
aggacctggg	aaaacgacaa	tgaatttgat	agcgacgtag	agagaagaaa	agatggaaac	1800
aacactcaag	aacaacagcg	tgtaaagaga	catgccattt	ctgtgtaa		1848

```
<210> 225
<211> 611
<212> PRT
<213> SHRIMP
```

<400> 225

Met 1	Glu	Tyr	Ile	Gly 5	Glu	Lys	Asn	Asn	Asn 10	Pro	Val	Ser	Asn	Glu 15	Ser
Val	Ser	Glu	Lys 20	Glu	Leu	Lys	Leu	Arg 25	Ser	Ser	Phe	Leu	Met 30	Ile	Gly
Lys	Lys	Thr 35	Ser	Lys	Tyr	Glu	Gln 40	Val	Met	Gly	Val	Tyr 45	Glu	Ala	Ile
Glu	Ser	Ile	Arg	Gln	Ser	Glu 55	Leu	Ser	Glu	Asp	Thr 60	Phe	Val	Val	His
Val 65	Lys	Lys	Asp	Lys	Gln 70	Leu	Lys	Phe	Arg	Leu 75	Lys	Arg	Leu	Gln	Glu 80
Leu	Val	Glu	Asp	Asp 85	Ser	Leu	Arg	Ile	Glu 90	Arg	Ile	Ser	Cys 95	Ala	Pro
Pro	Glu	Pro	Gly 100	His	Leu	Phe	Lys	Asp 105	Asp	Ala	Gly	His	Val 110	Thr	Asp
Glu	Glu	Trp 115	Leu	Ala	Thr	Gln	Glu 120	Glu	Asp	Val	Arg	Lys 125	Ile	Asn	Thr
Ile	Val	Lys	Glu	Lys	Leu	Lys 135	Arg	Lys	Asp	Lys	Asp 140	Phe	Lys	Phe	Ser
Gln 145	Leu	Tyr	Arg	Tyr	Met 150	Ser	Asn	Ser	Leu	Ser 155	Glu	Ala	Val	Glu	Lys 160
Lys	His	Asp	Cys	Met 165	Ile	Ile	Ser	Ser	Asp 170	Phe	Leu	Ile	Gly 175	Leu	Gly
Phe	Ser	Thr	Met 180	Asn	Val	Thr	His	Ala 185	Leu	Lys	Ser	Met 190	Glu	Arg	Thr
Met	Gln	Lys 195	His	Gly	Phe	Lys	Asp 200	Met	Met	Val	Pro	Leu 205	Val	Glu	Ile
Cys	His	Arg	Thr	His	Tyr	Lys 215	Gly	Glu	Tyr	Ile	Ala 220	Asn	Pro	Ile	Phe
Lys 225	Ser	His	Ser	Ser	His 230	Cys	Leu	Ile	Val	Pro 235	Leu	Phe	Met	Val	Ala 240
Gly	Val	Phe	Ala	Arg 245	Ser	Ala	His	Pro	Ser 250	Ala	Ala	Ser	Ile	Glu 255	Met
Tyr	Leu	Ser	Thr 260	Leu	Ala	Tyr	Ala	Val 265	Ile	Lys	Asp	Glu	Lys 270	Gln	Arg
Gln	Ile	Arg 275	Glu	Glu	Leu	Ala	Arg 280	Lys	Asn	Leu	Gln	Ile 285	Lys	Glu	Glu
Leu	Glu	Asn	Gln	Val	Glu	Lys 295	Thr	Thr	Lys	Val	Glu 300	Lys	Glu	Leu	Glu
Thr 305	Gln	Val	Val	Lys	Thr 310	Thr	Lys	Val	Glu	Lys 315	Glu	Leu	Glu	Thr	Gln 320
Val	Val	Lys	Lys	Glu	Glu	Tyr	Lys	Asn	Ser	Tyr	Ile	Glu	Thr	Glu	Glu

[illegible]

<210>	226
<211>	885
<212>	DNA
<213>	SHRIMP

<400> 226						
atggagtata	taggggaaca	aaaactgata	aatctcctgg	atgagacacc	tgaagaagac	60
gagttacagc	tacgctcttc	ttttctaattg	attggagaaa	aacaatatga	aaagtatgaa	120
gaagtaatatg	gtacatttga	agcagttgaa	actatccgaa	agagtgaatt	tagagacggc	180
gttttcattg	tacaattaaa	ggaaaaataaa	cacattactt	ttgaaggggg	actgaaagaa	240
ttaaagagagc	tcacagggyga	caattccctc	aagactagaat	cattattatc	ctctattaag	300
cctgagaaaag	gacacgttat	acttaaaaac	acgtctacaa	ctactgatga	cgagtggctt	360
gcttctcaag	acaaagacgt	acaggaagta	aataagctcg	taaaggaaaa	gacacggatg	420
ttgttttagtg	ggttttatatt	cagtccaact	tacagggtata	taacaaaaag	tcttcctcaa	480
ataccttttcg	gggaaaagta	acgtttttgtt	gttagcacag	attttttgat	tggtctcggg	540
tttagtgctg	atgatgtcat	ggaaaaatta	atagcgatag	agggaaatat	gcggaaaagt	600
ggtctgaaat	atacatgggt	gcctgtggcc	gaagtgtgcc	atctcaaaaa	atacaagggt	660
gatattgttg	taaaccccaat	tttcaagagt	tatcattccc	attgcctagt	tattccactg	720
gtgtatcttag	ggtacatggt	ttcccgtaat	gttcaacccc	catctctaga	agtggagacg	780
tatcttcttgg	cgtttagcgtt	tgcattttgat	ttgtacggca	gggaagaaat	gcgcaagtct	840
tgcatgcgat	tatgtgagga	tattttctaq	tgtaaaaqqq	gqtaa		885

WO 01/38351

316

PCT/US00/28888

<210> 227  
<211> 290  
<212> PRT  
<213> SHRIMP

<400> 227

Met	Glu	Tyr	Ile	Gly	Glu	Gln	Lys	Leu	Ile	Asn	Leu	Leu	Asp	Glu	Thr
1				5					10					15	
Pro	Glu	Glu	Asp	Glu	Leu	Gln	Leu	Arg	Ser	Ser	Phe	Leu	Met	Ile	Gly
			20					25					30		
Glu	Lys	Gln	Tyr	Glu	Lys	Tyr	Glu	Glu	Val	Met	Ser	Thr	Phe	Glu	Ala
		35					40					45			
Val	Glu	Thr	Ile	Arg	Lys	Ser	Glu	Phe	Arg	Asp	Gly	Val	Phe	Ile	Val
	50					55				60					
Gln	Leu	Lys	Glu	Asn	Lys	His	Ile	Thr	Phe	Glu	Gly	Gly	Leu	Lys	Glu
65					70					75				80	
Leu	Arg	Glu	Leu	Thr	Gly	Asp	Asn	Ser	Leu	Lys	Ile	Glu	Ser	Leu	Leu
				85				90						95	
Ser	Ser	Ile	Lys	Pro	Glu	Lys	Gly	His	Val	Ile	Leu	Lys	Asn	Thr	Ser
			100					105					110		
Thr	Thr	Thr	Asp	Asp	Glu	Trp	Leu	Ala	Ser	Gln	Asp	Lys	Asp	Val	Gln
		115					120						125		
Glu	Val	Asn	Lys	Leu	Val	Lys	Glu	Lys	Thr	Arg	Met	Leu	Phe	Arg	Gly
	130					135					140				
Phe	Tyr	Phe	Ser	Pro	Tyr	Tyr	Ile	Thr	Lys	Ser	Leu	Pro	Gln	Ile	Pro
145					150					155					160
Phe	Gly	Glu	Lys	Glu	Arg	Phe	Val	Val	Ser	Thr	Asp	Phe	Leu	Ile	Gly
				165					170					175	
Leu	Gly	Phe	Ser	Ala	Asp	Asp	Val	Met	Glu	Lys	Leu	Ile	Ala	Ile	Glu
			180					185					190		
Gly	Asn	Met	Arg	Lys	Ser	Gly	Leu	Lys	Tyr	Thr	Trp	Val	Pro	Val	Ala
	195						200					205			
Glu	Val	Cys	His	Leu	Lys	Lys	Tyr	Lys	Gly	Asp	Ile	Val	Val	Asn	Pro
	210					215					220				
Ile	Phe	Lys	Ser	Tyr	His	Ser	His	Cys	Leu	Val	Ile	Pro	Leu	Val	Tyr
225					230					235					240
Leu	Gly	Tyr	Met	Phe	Ser	Arg	Asn	Val	Gln	Pro	Pro	Ser	Leu	Glu	Val
				245					250					255	
Glu	Thr	Tyr	Leu	Leu	Ala	Phe	Ala	Ile	Asp	Leu	Tyr	Gly	Arg	Glu	Glu
			260					265					270		
Met	Arg	Lys	Ser	Cys	Met	Arg	Leu	Cys	Glu	Asp	Ile	Ser	Glu	Val	Lys
		275					280					285			
Arg	Gly														
	290														

<210> 228  
<211> 2769  
<212> DNA  
<213> SHRIMP

<400> 228

atggaggtcaa	tcaaggaaca	gcagcagcag	cagccaactg	tcactttctc	tgaggaacca	60
gatcaagttt	atgaatttga	ggatactact	actacttctg	ccaagaaacc	aactccttcc	120
aaggccaagt	ttgctgctgg	aagaagaatg	gtctctaagc	agaggaggaa	cactatcagg	180
agtccacaca	cggaaactgt	ggaagaagtt	gtaggagagg	aagaagaaca	gcagcagcag	240
actcctccag	aaatcacacc	cgccgagaag	aagcagcaat	ctcttcaaga	actggacgct	300
cttatgggca	aggtgccagc	tcactcttgac	gtgagtgttc	tggccaaatc	tggttgctgaa	360
tttctggaga	atgacgaaga	tgaggacgaa	gaattggaga	agaataagaa	ggcgacagaaa	420
tcggtactgt	tcaatagtgt	gatgaactct	ggacgcacag	aactgtcccc	ttccaccttt	480
tgtgatgggt	gtgtctcaaa	ggtcaagtcg	gcatttgaag	gaaaggatct	ggtttagcaac	540



```

attgtcaagg ttgagggaga agctgtgaag aaaactgcaa tcgccactga tactacaaaa 600
ctggccaatc tttccttgg ttgtatgaat ctccagttcc atgagcacgt gactattgaa 660
acactcaaca agaaggccct tgataaggga ggccctctct tcacattgaa gctctctgat 720
gctgtatatg tggacgagat ggatttggag aagaagaggc agatatttgg aagtaatgga 780
gataagtcac tgtttaagga acttggagga aactatattg atagtgcaat caagtcaact 840
ggatttggta tgtcaactcc cagttcttcg tctaccaaga aggcaggaac acatttcaag 900
accactaacc aaattgtgga agaaagtgtc actgaatcta tgagaaatgg atgttgttgc 960
ttcaagaacg acaagtgggt ggccaagagg gagagtaacc tcaagtcact gaacaatact 1020
gtatttggag aagaggatga cgaagagtca gcatatgcat acagcgacag tgaggatgaa 1080
gacgaagatg aaaacgagga ggaagtgtat tacgattaca acaatgaaac tattgaaagt 1140
tcagttggaa atgtcatcaa gaatcttatt agaaaaacta tcggtctcag tgatgtagag 1200
gaggaaaagg aagaggtgga gcagtcagag gaagaagaag aagactccga tgatgatgat 1260
gacgacgcct catcagtgtg ttcttcttcc tcttcgtctt catctgtcac tgtttagtagt 1320
gctgctgaag aagaagagga agaggatgag gaggataagg ataaggatac agcaacagta 1380
gtagaggatg aagatgacaa ggaaagtgtt atcagcagta gcagtgaaga ttcagaagaa 1440
gatgaggatg atgatggtgc aacttcccag tgttcagaag ttgtgtttgg ggatgtaaca 1500
gagtgcgaat tcgacgagag cgatgggaac cctctctacc tggcttcaga caacagcttc 1560
agaccatctg cttcagtgc ccaatatact caatctgagg aagagatgga tgtttctctc 1620
ctatcaaaga acagatcaac tcctgtctgt ctatctctat gcagacactc ctctggatgc 1680
atcaccaact cattcaacat gtccactatc ctcaagtctc ttaagtgtgt ccctgccgga 1740
actgaggctg ctgaggattg tgtgcacatc gactctacaa agaagaagga tgaggatgaa 1800
gatgaggagg atcagggact tgatctgcaa aacagccagt actactcagt acttgtcgat 1860
gtggacaatc ttattatctt ttctatgggt tccactacat acgaatcttc aatggtggag 1920
gtagattatg ataagagttt ctggtcctct ttcgacaagt ctgtgaagcc ttattgtgag 1980
agcaagaaga gtgctcttat taatgctcta tgtgaagata atgtgactgc caaggtatac 2040
gctactgtcc atactgtgc tattccattc tgtgaatcta tgcccattaa ccacattaat 2100
aacactactc cttatgggaag ctacaagacc ttcagaatta gtttgccagg gaacttttct 2160
gggcaacata acgacatcaa caacaattgg agatctgaca tgtactactaa gatggttgaa 2220
aatcttctaa agagggaagt tgttgagaac aagacacaca gccgcagata cgttcgcaac 2280
cttattgttg acggaggcgt gggagaaaat agtggcaact acttgaaggt gcacgagaat 2340
aatgaggata tttttggaag cattgaagcg aattcaatgt ctgccaaagac tgctgctgcc 2400
gcctttaaga atgttgccaa gaaatgtgac ctatccaga ctacaaccaa cgacatcttg 2460
actgggcctt tcaagcaata cctcattgat tacaagtata attcggccag aaagaatatt 2520
atcatggagc cttgtgaggg ggatgaaact acagcacatg agatgaagag ggctcaagac 2580
gcatacaaac aagcactcca cagggcaaag attacagcaa gctccatctc gctgaggggc 2640
atctggcacg agatgatcac cagggatatg aatacaactt acaatagcat gtttatgtat 2700
attcctgatt ttataaata tgttcaggta tcacctgtaa atgtatcacc attatatatg 2760
ttgattag 2769

```

<210> 229

<211> 922

<212> PRT

<213> SHRIMP

<400> 229

```

Met Glu Ser Ile Lys Glu Gln Gln Gln Gln Pro Thr Val Thr Phe
1 5 10 15
Ser Glu Glu Pro Asp Gln Val Tyr Glu Phe Glu Asp Thr Thr Thr
20 25 30
Ser Ala Lys Lys Pro Thr Pro Ser Lys Ala Lys Phe Ala Ala Gly Arg
35 40 45
Arg Met Val Ser Lys Gln Arg Arg Asn Thr Ile Arg Ser Pro His Thr
50 55 60
Glu Thr Val Glu Glu Val Val Gly Glu Glu Glu Glu Gln Gln Gln Gln
65 70 75 80
Thr Pro Pro Glu Ile Thr Pro Ala Glu Lys Lys Gln Gln Ser Leu Gln
85 90 95
Glu Leu Asp Ala Leu Met Gly Lys Val Pro Ala His Leu Asp Val Ser
100 105 110
Val Leu Ala Lys Ser Val Ala Glu Phe Leu Glu Asn Asp Glu Asp Glu
115 120 125

```

Asp	Glu	Glu	Leu	Glu	Lys	Asn	Lys	Lys	Ala	Gln	Lys	Ser	Val	Leu	Phe
130						135					140				
Asn	Ser	Val	Met	Asn	Ser	Gly	Arg	Thr	Glu	Leu	Ser	Pro	Ser	Thr	Phe
145					150					155					160
Cys	Asp	Gly	Cys	Val	Ser	Lys	Val	Lys	Ser	Ala	Phe	Glu	Gly	Lys	Asp
				165					170					175	
Leu	Val	Ser	Asn	Ile	Val	Lys	Val	Glu	Gly	Glu	Ala	Val	Lys	Lys	Thr
			180					185					190		
Ala	Ile	Ala	Thr	Asp	Thr	Thr	Lys	Leu	Ala	Asn	Leu	Phe	Leu	Gly	Cys
		195					200					205			
Met	Asn	Leu	Gln	Phe	His	Glu	His	Val	Thr	Ile	Glu	Thr	Leu	Asn	Lys
	210					215					220				
Lys	Ala	Leu	Asp	Lys	Gly	Gly	Pro	Leu	Phe	Thr	Leu	Lys	Leu	Ser	Asp
225				230						235					240
Ala	Val	Tyr	Val	Asp	Glu	Met	Asp	Leu	Glu	Lys	Lys	Arg	Gln	Ile	Phe
				245					250					255	
Gly	Ser	Asn	Gly	Asp	Lys	Ser	Leu	Phe	Lys	Glu	Leu	Gly	Gly	Asn	Tyr
			260					265					270		
Ile	Asp	Ser	Ala	Ile	Lys	Ser	Thr	Gly	Leu	Val	Met	Ser	Thr	Pro	Ser
	275						280					285			
Ser	Ser	Ser	Thr	Lys	Lys	Ala	Gly	Thr	His	Phe	Lys	Thr	Thr	Asn	Gln
	290					295					300				
Ile	Val	Glu	Glu	Ser	Val	Thr	Glu	Ser	Met	Arg	Asn	Gly	Cys	Cys	Cys
305					310					315					320
Phe	Lys	Asn	Asp	Lys	Trp	Leu	Ala	Lys	Arg	Glu	Ser	Asn	Leu	Lys	Ser
				325					330					335	
Leu	Asn	Asn	Thr	Val	Phe	Gly	Glu	Glu	Asp	Asp	Glu	Lys	Ser	Ala	Tyr
			340					345					350		
Ala	Tyr	Ser	Asp	Ser	Glu	Asp	Glu	Asp	Glu	Asp	Glu	Asn	Glu	Glu	Glu
		355					360					365			
Val	Asp	Tyr	Asp	Tyr	Asn	Asn	Glu	Thr	Ile	Glu	Ser	Ser	Val	Gly	Asn
	370					375					380				
Val	Ile	Lys	Asn	Leu	Ile	Arg	Lys	Thr	Ile	Gly	Leu	Ser	Asp	Val	Glu
385					390					395					400
Glu	Glu	Lys	Glu	Glu	Gly	Glu	Gln	Ser	Glu	Glu	Glu	Glu	Glu	Asp	Ser
				405					410					415	
Asp	Asp	Asp	Asp	Asp	Asp	Ala	Ser	Ser	Val	Cys	Ser	Ser	Ser	Ser	Ser
			420					425					430		
Ser	Ser	Ser	Val	Thr	Val	Val	Ala	Ala	Ala	Glu	Glu	Glu	Glu	Glu	Glu
		435					440					445			
Asp	Glu	Glu	Asp	Lys	Asp	Lys	Asp	Thr	Ala	Thr	Val	Val	Glu	Asp	Glu
	450					455					460				
Asp	Asp	Lys	Glu	Ser	Val	Ile	Ser	Ser	Ser	Ser	Glu	Asp	Ser	Glu	Glu
465					470					475					480
Asp	Glu	Asp	Asp	Asp	Gly	Ala	Thr	Ser	Gln	Cys	Ser	Glu	Val	Val	Phe
				485					490					495	
Gly	Asp	Val	Thr	Glu	Cys	Glu	Phe	Asp	Glu	Ser	Asp	Gly	Asn	Pro	Leu
			500					505					510		
Tyr	Leu	Ala	Ser	Asp	Asn	Ser	Phe	Arg	Pro	Ser	Ala	Ser	Val	Thr	Lys
		515					520					525			
Tyr	Pro	Gln	Ser	Glu	Glu	Glu	Met	Asp	Val	Ser	Leu	Leu	Ser	Lys	Asn
	530					535					540				
Arg	Ser	Thr	Pro	Val	Cys	Leu	Ser	Leu	Cys	Arg	His	Ser	Ser	Gly	Cys
545					550					555					560
Ile	Thr	Asn	Ser	Phe	Asn	Met	Ser	Thr	Ile	Leu	Lys	Ser	Leu	Lys	Leu
				565					570					575	
Phe	Pro	Ala	Gly	Thr	Glu	Ala	Ala	Glu	Asp	Cys	Val	His	Ile	Glu	Ser
			580					585					590		
Thr	Lys	Lys	Lys	Asp	Glu	Asp	Glu	Asp	Glu	Glu	Asp	Gln	Gly	Leu	Asp
		595					600					605			
Leu	Gln	Asn	Ser	Gln	Tyr	Tyr	Ser	Val	Leu	Val	Asp	Val	Asp	Asn	Leu

WO 01/38351

319

PCT/US00/28888

610	Ile	Ile	Phe	Ser	Met	Gly	615	Ser	Thr	Thr	Tyr	Glu	620	Ser	Ser	Met	Val	Glu
625						630						635					640	
Val	Asp	Tyr	Asp	Lys	Ser	Phe	Trp	Ser	Ser	Ser	Phe	Asp	Lys	Ser	Val	Lys		
				645							650					655		
Pro	Tyr	Cys	Glu	Ser	Lys	Lys	Ser	Ala	Leu	Ile	Asn	Ala	Leu	Cys	Glu			
			660					665					670					
Asp	Asn	Val	Thr	Ala	Lys	Val	Tyr	Ala	Thr	Val	His	Thr	Leu	Ala	Ile			
		675					680					685						
Pro	Phe	Cys	Glu	Ser	Met	Pro	Ile	Asn	His	Ile	Asn	Asn	Thr	Thr	Pro			
		690					695					700						
Tyr	Gly	Ser	Tyr	Lys	Thr	Phe	Arg	Ile	Ser	Leu	Pro	Gly	Asn	Phe	Ser			
705					710					715					720			
Gly	Gln	His	Asn	Asp	Ile	Asn	Asn	Asn	Trp	Arg	Ser	Asp	Met	Tyr	Thr			
				725					730					735				
Lys	Met	Val	Glu	Asn	Leu	Leu	Lys	Arg	Glu	Val	Val	Glu	Asn	Lys	Thr			
			740					745					750					
His	Ser	Arg	Arg	Tyr	Val	Arg	Asn	Leu	Ile	Val	Asp	Gly	Gly	Val	Gly			
		755					760					765						
Glu	Asn	Ser	Gly	Asn	Tyr	Leu	Lys	Val	His	Glu	Asn	Asn	Glu	Asp	Ile			
		770					775					780						
Phe	Gly	Ser	Ile	Glu	Ala	Asn	Ser	Met	Ser	Ala	Lys	Thr	Ala	Ala	Ala			
785					790					795					800			
Ala	Phe	Lys	Asn	Val	Ala	Lys	Lys	Cys	Asp	Leu	Ile	Gln	Thr	Thr	Thr			
				805					810					815				
Asn	Asp	Ile	Leu	Thr	Gly	Pro	Phe	Lys	Gln	Tyr	Leu	Ile	Asp	Tyr	Lys			
			820					825					830					
Tyr	Asn	Ser	Ala	Arg	Lys	Asn	Ile	Ile	Met	Glu	Pro	Cys	Glu	Gly	Asp			
		835					840					845						
Glu	Thr	Thr	Ala	His	Glu	Met	Lys	Arg	Ala	Gln	Asp	Ala	Tyr	Lys	Gln			
		850				855					860							
Ala	Leu	His	Arg	Ala	Lys	Ile	Thr	Ala	Ser	Ser	Ile	Ser	Leu	Arg	Gly			
865					870					875					880			
Ile	Trp	His	Glu	Met	Ile	Thr	Arg	Asp	Met	Asn	Thr	Thr	Tyr	Asn	Ser			
				885					890					895				
Met	Phe	Met	Tyr	Ile	Pro	Asp	Phe	Tyr	Lys	Tyr	Val	Gln	Val	Ser	Pro			
			900					905					910					
Val	Asn	Val	Ser	Pro	Leu	Tyr	Met	Leu	Asp									
		915					920											

<210> 230  
 <211> 846  
 <212> DNA  
 <213> SHRIMP

<400> 230  
 atgtctcctg ttatctoccca acaatcttcc ccttctgcta ccagcacggc tgcagccaga 60  
 ataattagta ctgccaaact tagagttcta ggtgtcaaaa ataaggaaga gaaagatgag 120  
 gaagaagaac aacaggaagt tgaaccagaa attattgaac cagctaccga ttttgagata 180  
 ccattttctc cagcactcac aatctgcata tacatcaatg cgaatcgcat acacattaat 240  
 tcaaaggggtg tgtgcttaaa taggaaaaag ataaaacctt cttcaacaat aaacaaaaaac 300  
 caggacgtcc ctccagaact agcaaatgctg tccagttatc ttgtacaaac tgaacacgta 360  
 accgacaagt tccttttcac ccactgttct atatgcaact ataacgtgaa cgacggggaa 420  
 tacaatcggt ctctaagcac aacaagaaat ggagatcagc ccttgatgag aaagtcgggtc 480  
 agatatgttc ccttaaacga agataatgtg gtagtccaga aaggaacata ttatgggact 540  
 actttttatac cagaaaagac gggagaagaaga attttgtggt tctctcatta caagaagtct 600  
 cctcgtccaa ttacagctaa attatgttgt cttctggaaa ctataaaactc gttcaacggc 660  
 agctgttctt cttctctctc tgcgtcctcc tccagcaacg ccccagggcc tattgaagaa 720  
 ttccaagtgt cttcatccat atttttcaag aaagaagagt gttgtcccct gcaaatgaag 780  
 tgggttgaac aaaatgagct ggatgcagaa tcacctgttc ttgtacttct aatgttagcg 840

WO 01/38351

320

PCT/US00/28888

ctatga

846

<210> 231  
 <211> 281  
 <212> PRT  
 <213> SHRIMP

<400> 231

Met	Ser	Pro	Val	Ile	Ser	Gln	Gln	Ser	Ser	Pro	Ser	Ala	Thr	Ser	Thr
1				5					10					15	
Ala	Ala	Ala	Arg	Ile	Ile	Ser	Thr	Ala	Asn	Leu	Arg	Val	Leu	Gly	Val
			20					25					30		
Lys	Asn	Lys	Glu	Glu	Lys	Asp	Glu	Glu	Glu	Glu	Gln	Gln	Glu	Val	Glu
		35					40					45			
Pro	Glu	Ile	Ile	Glu	Pro	Ala	Thr	Asp	Phe	Glu	Ile	Pro	Phe	Ser	Pro
		50					55				60				
Ala	Leu	Thr	Ile	Cys	Ile	Tyr	Ile	Asn	Ala	Asn	Arg	Ile	His	Ile	Asn
65					70					75					80
Ser	Lys	Gly	Val	Cys	Leu	Asn	Arg	Lys	Lys	Ile	Lys	Pro	Thr	Ser	Thr
				85					90					95	
Ile	Asn	Lys	Asn	Gln	Asp	Val	Pro	Pro	Glu	Leu	Ala	Asn	Ala	Ser	Ser
			100					105					110		
Tyr	Leu	Val	Gln	Thr	Glu	His	Val	Thr	Asp	Lys	Phe	Leu	Ser	Ser	His
		115					120					125			
Cys	Ser	Ile	Cys	Asn	Tyr	Asn	Val	Asn	Asp	Gly	Glu	Tyr	Lys	Ser	Ala
		130				135					140				
Leu	Ser	Thr	Thr	Arg	Asn	Gly	Asp	Gln	Pro	Leu	Met	Arg	Lys	Ser	Val
145					150					155					160
Arg	Tyr	Val	Pro	Leu	Asn	Glu	Asp	Asn	Val	Val	Val	Gln	Lys	Gly	Thr
				165					170					175	
Tyr	Tyr	Gly	Thr	Thr	Phe	Ile	Pro	Glu	Lys	Thr	Gly	Arg	Arg	Ile	Leu
			180					185					190		
Trp	Phe	Ser	His	Tyr	Lys	Lys	Ser	Pro	Arg	Pro	Ile	Thr	Ala	Lys	Leu
		195					200					205			
Cys	Cys	Leu	Leu	Glu	Thr	Ile	Asn	Ser	Phe	Asn	Gly	Ser	Cys	Ser	Ser
	210					215					220				
Ser	Ser	Ser	Ala	Ser	Ser	Ser	Ser	Asn	Ala	Pro	Gly	Pro	Ile	Glu	Glu
225					230					235					240
Phe	Gln	Val	Ser	Ser	Ser	Ile	Phe	Phe	Lys	Lys	Glu	Glu	Cys	Cys	Pro
				245					250				255		
Leu	Gln	Met	Lys	Trp	Val	Glu	Gln	Asn	Glu	Leu	Asp	Ala	Glu	Ser	Pro
			260					265					270		
Val	Leu	Val	Leu	Leu	Met	Leu	Ala	Leu							
		275					280								

<210> 232  
 <211> 1470  
 <212> DNA  
 <213> SHRIMP

<400> 232

atggaatctg	taagagacgt	taaattttac	acttttatga	atgtattggc	agagaaggcg	60
aaaaagatac	agcgccataa	caaggataaa	ggttggagga	cttcaataaa	cgctgaaatt	120
ggatatggag	gagcaagact	catggatgtg	cgcttcacgg	gcaggaaatc	aatggacgaa	180
ttggcaagat	gtctctataa	ctgcgatgga	gaatatacaa	ctcttagact	cgtgggttca	240
agtgcgggaa	atattatagt	atattcttta	gcctttataa	tgggaatcag	aggcgagtgt	300
tgtgggttta	atgtcaataa	tcgcctacga	atgggaaaaa	taatagacag	agagttgttt	360
tataaaataa	caggattaaa	tttccctgaa	actgtaaaaa	gcacatgtga	tggagtacga	420
gccatctgtg	atttgttctt	ggaggtggct	gctcttcaag	aacaccccg	ctggcatgaa	480
actaaagaag	tagggaaaaa	acaacaacaa	catttcaacg	aattcggggtc	ccagtaccct	540

```

ggcacaaaaat tcaataagcgc gcacaaaacta tccactaaaaa taatttcagca aatgtttttca 600
gaagagaaga ctatggaaca agttctagcc tttagtgaag gaactgccgc tagtggattt 660
tcagattttgt acgtagaagc acctatacaa tacgtggtta atatgtatag agctatcagt 720
aatatggaag gaagagtagg tgccatgtat aatttgtcga gagtttttaat actcttatgt 780
tctagggtggg aaaaagaaacc aggggtataaa aatgactttt atagtaagtg tgagatgtac 840
ataggctcca aaaaaattgt tgatgacgag agctttattt tcaccgatct cataactggg 900
gatttagttc ctctagttcg attggcccca tcaaacgaag atatacagag agatgtttata 960
aggttcaatg attctactga tatattaatg gacagtattg acgtgaggga tgttgttttg 1020
cctgtacttt ctaaaataat atggcaaaac gtttctgcac gtctaaaatt gcgcaacaat 1080
aaatccctct ccaaactagc caaatggaaa tggaatggta tgggtgtcaac acacgacaac 1140
tttgattcga atgattatgt cattgaacat aaacgtcaac ttgcagcgga tataatgtcc 1200
gatagtctat caaaaaatca ccttccaaat tttagtaaaa caataacgga gtacgatgaa 1260
aaggaaaata aaacaacacc cttaatatgt tggaattata tttttgaatt gtcccctatg 1320
gggaaacact tgttccttt agaggaagtt tgcggtattc atgaggcgag tttacctctc 1380
ataacacccat ggcaattaaa ggttgtccag aaaaagagag gaaggcagat ggtcatctac 1440
ggtcctagaa aacgacctcg cacacagtaa 1470

```

&lt;210&gt; 233

&lt;211&gt; 487

&lt;212&gt; PRT

&lt;213&gt; SHRIMP

&lt;400&gt; 233

```

Met Glu Ser Val Arg Asp Val Lys Phe Tyr Thr Phe Met Asn Val Leu
 1          5          10          15
Ala Glu Lys Ala Lys Lys Ile Gln Arg Leu Asn Lys Asp Lys Gly Trp
 20          25          30
Arg Thr Ser Ile Asn Ala Glu Ile Gly Tyr Gly Gly Ala Arg Leu Met
 35          40          45
Asp Val Arg Phe Thr Gly Arg Lys Ser Met Asp Glu Leu Ala Arg Cys
 50          55          60
Leu Tyr Asn Cys Asp Gly Glu Tyr Thr Thr Leu Arg Leu Val Gly Ser
 65          70          75          80
Ser Ala Gly Asn Ile Ile Val Tyr Ser Leu Ala Phe Ile Met Gly Ile
 85          90          95
Arg Gly Glu Cys Cys Gly Phe Asn Val Asn Asn Arg Leu Arg Met Gly
100          105          110
Lys Ile Ile Asp Arg Glu Leu Phe Tyr Lys Ile Thr Gln Phe Pro Glu
115          120          125
Thr Val Lys Cys Thr Cys Asp Gly Val Arg Ala Ile Cys Asp Leu Phe
130          135          140
Leu Glu Val Ala Ala Leu Gln Glu His Pro Ala Trp His Glu Thr Lys
145          150          155          160
Glu Val Gly Lys Lys Gln Gln Gln His Phe Asn Glu Phe Gly Ser Gln
165          170          175
Tyr Pro Gly Thr Lys Phe Asn Lys Arg His Lys Leu Ser Thr Lys Ile
180          185          190
Ile Gln Gln Met Phe Ser Glu Glu Lys Thr Met Glu Gln Val Leu Ala
195          200          205
Phe Ser Glu Gly Thr Ala Ala Ser Gly Phe Ser Asp Leu Tyr Val Glu
210          215          220
Ala Pro Ile Gln Tyr Val Val Asn Met Tyr Arg Ala Ile Ser Asn Met
225          230          235          240
Glu Gly Arg Val Gly Ala Met Tyr Asn Leu Ser Arg Val Leu Ile Leu
245          250          255
Leu Cys Ser Arg Trp Glu Lys Lys Pro Gly Tyr Lys Asn Asp Phe Tyr
260          265          270
Ser Lys Cys Glu Met Tyr Ile Gly Ser Lys Lys Ile Val Asp Asp Glu
275          280          285
Ser Phe Ile Phe Thr Asp Leu Ile Thr Gly Asp Leu Val Pro Leu Val
290          295          300

```

WO 01/38351

322

PCT/US00/28888

Arg Leu Ala Pro Ser Asn Glu Asp Ile Gln Arg Asp Val Ile Arg Phe  
 305 310 315 320  
 Asn Asp Ser Thr Asp Ile Leu Met Asp Ser Ile Asp Val Arg Asp Val  
 325 330 335  
 Val Leu Pro Val Leu Ser Lys Ile Ile Trp Gln Asn Val Ser Ala Arg  
 340 345 350  
 Leu Lys Leu Arg Asn Asn Lys Ser Leu Ser Lys Leu Ala Lys Trp Lys  
 355 360 365  
 Trp Asn Gly Met Val Ser Thr His Asp Asn Phe Asp Ser Asn Asp Tyr  
 370 375 380  
 Val Ile Glu His Lys Arg Gln Leu Ala Ala Asp Ile Met Ser Asp Ser  
 385 390 395 400  
 Leu Ser Lys Asn His Leu Pro Asn Phe Ser Lys Thr Ile Thr Glu Tyr  
 405 410 415  
 Asp Glu Lys Glu Asn Lys Thr Thr Pro Leu Ile Cys Trp Asn Tyr Ile  
 420 425 430  
 Phe Glu Leu Ser Pro Met Gly Lys His Leu Phe Pro Leu Glu Glu Val  
 435 440 445  
 Cys Gly Phe Tyr Glu Ala Ser Leu Pro Leu Ile Thr Pro Trp Gln Leu  
 450 455 460  
 Lys Val Val Gln Lys Lys Arg Gly Arg Gln Met Val Ile Tyr Gly Pro  
 465 470 475 480  
 Arg Lys Arg Pro Arg Thr Gln  
 485

<210> 234  
 <211> 4698  
 <212> DNA  
 <213> SHRIMP

<400> 234  
 atgtcaacaa cgcaaacgca aacaatagaa cggcccctgc ctggtaagaa caacgaagac 60  
 aatagccgtt tagcttgccct attggctgaa gggctacagc aacaacaaca acaacaagat 120  
 ggcgatagcg aaatttctct accgcttggt aatgcgggaa cttttgcgtg ttatgattct 180  
 acccttgcaa acctcaccga ggggcgttta ggaagtgaag cagaaaatgc aaaaataagg 240  
 gtaaaaatac acccgctctgt gtttattata gagacgaata aagagatgac tattgaagaa 300  
 atatctacaa aatctctaaa tgcattagtg gaaaaagag cacgagaagc acgacgattt 360  
 tcatctctga cagaacaaaa atttctctga ggaggaggag gatgttattc tcgtaaaaaat 420  
 gagaggttta tcgaaggcga aattaacaac atcaagctga acatggaaga aactgcttcg 480  
 tctttagaga gattggcagg gcttttgccct gttgtcatta atattaagga ctggacaatg 540  
 catgatgaga aagaaatacg gctagatctt aaaggaaacg acggcatgga agaactgttc 600  
 aatatatccc atctgaatca agaagaatgg gaaatggaaa gactttcttc ctcaatagta 660  
 ttgaaagacg cctatggtgt gttttatgcc catcacggca tacttgatat tgttcttacg 720  
 acctctagat tcacagggaa attattacaa caccctgtga tatttcgtct aatggatgtg 780  
 aaggtgtgga taaacacccc cctccagata gcatttcctg acaccagtaa gaaccctaat 840  
 gcaaaagaaa ttctctatca acatccctct cttacaaggc tacgtgattt aaatgacatg 900  
 gcttcaaaact ccaagtcagt ttcttccatt atcataccag agttgtcaaa atttaactcg 960  
 actgaatttg gtatgcacta ctttacggcc cagtgccttt ttggaaaaaa taccaattct 1020  
 ttaaaggatt tgggtacacg gtattatcaa ctatcattta aaaataaacc tcaacccaaa 1080  
 ctttatgaac caacagcaac agcaacagca gcctcctcct cctcctcaac cgctcattg 1140  
 acgaccgaac aaaaggaaaa aatcgcacaa tcgattttat cgtctaaagg aaaatcttta 1200  
 ggagatgtat cttccaccct ttctaaggaa tacgatgaaa atagaaaacg tacaagagg 1260  
 caaaaaactt ctaccgatac caacattgtc ccttcagggg cacctacttc catttccatg 1320  
 aaaaacccag ttacttgctt ctttgggcct caatacactt caatcatgga ctgtatatcg 1380  
 gagaaaacag actggataga aatgcacctt tttttaacct ctcttaacga cgctgaacat 1440  
 aataaaacac tcgttgctga cagaaaaagt aacgtgtcgg aaatacacga ttctggaagg 1500  
 tttcttatcat tcggtcaaaa taatactaca gcctttatag catagtgtga tattcctacg 1560  
 ttaaaattaa ttctgagaga cgattctggg gaatcttctg ctgctataat agcatctctg 1620  
 atatattata ataatgtaaa tctagaaggg aggggaattct caaatgtctc tgatgctgtt 1680  
 gtgggacttt tctcgggcgg atctgctatt acggtaggag atattgctcg agagattgca 1740

```

tccatataca atattggaag agaatctaac tgcgattcca tactattccc aggggagcct 1800
atactggccg gaagaagaag ttatggacgg cagtacagat ggtacgatcc cattaactgc 1860
gtggtgggtc tttatcgggtc gtgttttagaa acaatgacaa gaaatattat gagaggacag 1920
ccagtgaag tggtgagac cgcttggtg tacatgcacc agcagggtact tcaagttggt 1980
cttcttccat tttttgattg cgtttttgaaa ctgggctgtaa gggctgtaaa ggaggctaga 2040
caactaacag attatatagt acgtgaagtg ctactaaaat atacagccga tcctgatcaa 2100
cacaaatttt tattattcaa gaagcccgtg atggatctaa ttgcaaagat tgttacacat 2160
tatgcagtta ttcatcggc agccgataat ggaggcgtgt gccttgcttt ccctagagat 2220
cctcccttta ttgtagaaaa tgatacgtct ctcagatact atacactgac cgatacacct 2280
cagagtattt tgaatggaga taatgtagcc gaaaacttaa agtctgccac ttcagtcgct 2340
tcttctccat catcctcctc cagataattca tcagaaaccc ctattagggt agttaatctt 2400
cctgtaccaa caggacggtt tttaaagatg aataaggatc tggaaacttt cattaatgta 2460
cctctaatat cctctaagga acaaaaacaa caacagcagc aaacaactgc cactgctcca 2520
tttcttctg aaaccatttc aaaatcggtt ttaaattatg taccacctaa atctctaaca 2580
agaaatgtga catatggtca gaatatcgca gaagatggat tccttggatt gaaaaataaa 2640
ggtgaattgg tgtcgtactt taaagtgtga aagaatactg aacgcgatga tggtataaag 2700
gatatggaaa ttggtgatat taataacccat caagacgaca ctggatccct ttcttcttct 2760
tcttcttcat tcgttgatgg tgtacggaca agtttttcgg ttgatgggaa aattgaacat 2820
gttagcgcat tcctccctgg tacgacttct cagcccacaa atcttctctgt tcatgcatct 2880
aaacaagtca aatattcggg aaaggaattg ggcatggaaa ctgtattctt tgaaccctt 2940
ctctcgtctg ctgtgcttta tgaagcctca aaaactaaat cgacacaaca ttttaagcca 3000
atgaggatct ataaggaatg tgtagtcct ctatctacag ggaggatcga tattttcccc 3060
agtaaagtgg gcacagtggc aggaactgga ttcgaattta tttggaaagt tcttcagtat 3120
gatactggat taccacaac tctagaaagg ttgtcaccce agattccttc tgtaccaatt 3180
tcaggcgaag actcaaaaat ggaagtaatt gcagaatctg gtaaagggtg acaaaacatc 3240
attgctattg cagccgacca actcagaggc agtaataata tagtaggtgg tggaacaaga 3300
agagcgatcc aacaacaaca acagcagcaa caacaagaaa cacaagcggt tgttcagtt 3360
aatgtgcctg cccgtttcga gcctacgttc acagagatcg aactatttct tcagaacaaa 3420
tttaggaatg ttattgctac tattatatct agaattgatg tgcttgctcag caacgaagag 3480
atgaaaatta ttaagaagat gtgcgagcac gtgtctcata ttatggtcga tggattatat 3540
gttgacttag atcctagaaa ggccattgaa gaaatactgg aaaggataac agctgaacaa 3600
aatgggatta ctattgatac tggaaatgaa gggtatggat ctttacgcta tgcttcttcg 3660
gggagacttt ttataaatga tgaagctagt gaagaagctg cagcggccat cggcgggtgga 3720
ggtgctcttg gaactggtag aagggtgcct gtagaattac gttcaatatt ggacaaatta 3780
aataccattg gaagtactac tcaacaacaa caacagcagc agcgtcaaca gaggcaggct 3840
aataataaca atactgttcc cgaggatata aaggttcaca atgaacaaat gcaaaaaatt 3900
agagactctt ctctattcac atctaaaact ttgaattaca ttagggatga tggaagaaag 3960
gaccgtataa agacaaacat ttcagaaaca cttaaagaaat acagtagaat cccatcatc 4020
tttattgctt caaaggcaca gaaacctatc ccatggaaac acaccaagga caatatcaac 4080
ctcaacaaga ttccagaaga cttgaacttt tctcctgcac agaattctgt tggtccagtt 4140
aatccgcgcc atatecttac tgacatgcaa tggctcaact gtatttccat catcgaaact 4200
gcaacaagag attcggctat cgtcatgcaa agtttccaag agcaagcgga caagactaca 4260
actcaattgg aggaactttt atctcaatgg ataataattg tatctcaagt aactgatgaa 4320
aaatccccag cttatgtttc aagtgtaaaa ttggaatggt taaataatga agctagccgt 4380
attgctgcta tacgagagaa ttcagaaaaa tcaaaaattg taatgggtgt tcaggggaag 4440
attgttaaca ttgatgagct aggtatagtg gctgttgcca ggtcaatagt agacgttgat 4500
ttttatataa aaatgcccac tgtgtgggct tctagggatt ggaagaacct gatttattat 4560
gctgttaata tagcagcaac tccccttatt aataatatat cgaggggtat catggctgct 4620
tcacagacat ctgttttata tgactcttct cttgctctta ttgctgccga gcaagctaca 4680
agaaatataa ccatgtga                                     4698

```

<210> 235  
 <211> 1551  
 <212> PRT  
 <213> SHRIMP

<400> 235  
 Met Ser Thr Thr Gln Thr Gln Thr Ile Glu Arg Pro Leu Pro Gly Lys  
 1 5 10 15  
 Asn Asn Glu Asp Asn Ser Arg Leu Ala Cys Leu Leu Ala Glu Gly Leu  
 20 25 30

Gln	Gln	Gln	Gln	Gln	Gln	Gln	Asp	Gly	Asp	Ser	Glu	Ile	Ser	Leu	Pro
		35					40					45			
Leu	Val	Asn	Ala	Gly	Thr	Phe	Ala	Cys	Tyr	Asp	Ser	Thr	Leu	Ala	Asn
	50					55					60				
Leu	Thr	Glu	Gly	Arg	Leu	Gly	Ser	Glu	Thr	Glu	Asn	Ala	Lys	Ile	Arg
65					70					75					80
Val	Lys	Ile	His	Pro	Ser	Val	Phe	Ile	Ile	Glu	Thr	Asn	Lys	Glu	Met
			85						90					95	
Thr	Ile	Glu	Glu	Ile	Ser	Thr	Lys	Ser	Leu	Asn	Ala	Leu	Val	Glu	Lys
			100					105					110		
Arg	Ala	Arg	Glu	Ala	Arg	Arg	Phe	Ser	Ser	Leu	Thr	Glu	Gln	Lys	Phe
		115					120					125			
Pro	Arg	Gly	Gly	Gly	Gly	Cys	Arg	Lys	Asn	Glu	Arg	Phe	Ile	Glu	Gly
	130					135				140					
Glu	Ile	Asn	Asn	Ile	Lys	Leu	Asn	Met	Glu	Glu	Thr	Ala	Ser	Ser	Leu
145					150					155					160
Glu	Arg	Leu	Ala	Gly	Leu	Leu	Pro	Val	Val	Ile	Asn	Ile	Lys	Asp	Trp
				165					170					175	
Thr	Met	His	Asp	Glu	Lys	Glu	Ile	Arg	Leu	Asp	Leu	Lys	Gly	Asn	Asp
			180					185					190		
Gly	Met	Glu	Glu	Leu	Val	Asn	Ile	Ser	His	Leu	Asn	Gln	Glu	Glu	Trp
		195				200						205			
Glu	Met	Glu	Arg	Leu	Ser	Ser	Ser	Ile	Val	Leu	Lys	Asp	Ala	Tyr	Gly
	210					215				220					
Val	Phe	Tyr	Ala	His	His	Gly	Ile	Leu	Asp	Ile	Val	Leu	Thr	Thr	Ser
225					230					235					240
Arg	Phe	Thr	Gly	Lys	Leu	Leu	Gln	His	Pro	Val	Ile	Phe	Arg	Leu	Met
				245					250					255	
Asp	Val	Lys	Val	Trp	Ile	Asn	Thr	Pro	Leu	Gln	Ile	Ala	Phe	Pro	Asp
			260					265					270		
Thr	Ser	Lys	Asn	Pro	Asn	Ala	Lys	Lys	Ile	Leu	Tyr	Gln	His	Pro	Ser
		275					280					285			
Leu	Thr	Arg	Leu	Arg	Asp	Leu	Asn	Asp	Met	Asn	Ser	Lys	Ser	Val	Ser
	290					295					300				
Ser	Ile	Ile	Ile	Pro	Glu	Leu	Ser	Lys	Phe	Asn	Ser	Thr	Glu	Phe	Gly
305					310					315					320
Met	His	Tyr	Phe	Thr	Ala	Gln	Cys	Phe	Phe	Gly	Lys	Asn	Thr	Asn	Ser
				325					330					335	
Leu	Lys	Asp	Leu	Val	Thr	Arg	Tyr	Tyr	Gln	Leu	Ser	Phe	Lys	Asn	Lys
			340					345					350		
Pro	Gln	Pro	Lys	Leu	Tyr	Glu	Pro	Thr	Ala	Thr	Ala	Thr	Ala	Ala	Ser
		355					360					365			
Ser	Ser	Ser	Ser	Thr	Ala	Ser	Leu	Thr	Thr	Glu	Gln	Lys	Glu	Lys	Ile
	370					375					380				
Ala	Gln	Ser	Ile	Leu	Ser	Ser	Lys	Gly	Lys	Ser	Leu	Gly	Asp	Val	Ser
385					390					395					400
Ser	Thr	Leu	Ser	Lys	Glu	Tyr	Asp	Glu	Asn	Arg	Lys	Arg	Thr	Lys	Arg
				405					410					415	
Gln	Lys	Thr	Ser	Thr	Asp	Thr	Asn	Ile	Val	Pro	Ser	Gly	Ala	Pro	Thr
			420					425					430		
Ser	Ile	Ser	Met	Lys	Asn	Pro	Val	Thr	Cys	Phe	Phe	Gly	Pro	Gln	Tyr
		435					440					445			
Thr	Ser	Ile	Met	Asp	Cys	Ile	Ser	Glu	Lys	Thr	Asp	Trp	Ile	Glu	Met
	450					455					460				
His	Leu	Phe	Leu	Thr	Ser	Leu	Asn	Asp	Ala	Glu	His	Asn	Lys	Thr	Leu
465					470					475					480
Val	Val	Asp	Arg	Lys	Ser	Asn	Val	Ser	Glu	Ile	His	Asp	Ser	Gly	Arg
				485					490					495	
Phe	Leu	Thr	Phe	Gly	Gln	Asn	Asn	Thr	Thr	Ala	Phe	Ile	Pro	Asp	Val
			500					505					510		
Asp	Ile	Pro	Thr	Leu	Lys	Leu	Ile	Leu	Arg	Asp	Asp	Ser	Gly	Glu	Ser



```

      515      520      525
Ser Ala Ala Ile Ile Ala Ser Leu Ile Tyr Tyr Asn Asn Val Asn Leu
530      535      540
Glu Gly Arg Glu Phe Ser Asn Val Ser Asp Ala Val Val Gly Leu Phe
545      550      555
Ser Gly Gly Ser Ala Ile Thr Val Gly Asp Ile Ala Arg Glu Ile Ala
565      570      575
Ser Ile Tyr Asn Ile Gly Arg Glu Ser Asn Cys Asp Ser Ile Leu Phe
580      585      590
Pro Gly Glu Pro Ile Leu Ala Gly Arg Arg Ser Tyr Gly Arg Gln Tyr
595      600      605
Arg Trp Tyr Asp Pro Ile Asn Cys Val Val Gly Arg Ser Cys Leu Glu
610      615      620
Thr Met Thr Arg Asn Ile Met Arg Gly Gln Pro Val Lys Val Asp Glu
625      630      635
Thr Ala Trp Met Tyr Met His Gln Gln Val Leu Gln Val Val Leu Leu
645      650      655
Pro Phe Phe Asp Cys Val Leu Lys Ser Gly Val Trp Ala Val Lys Glu
660      665      670
Ala Arg Gln Leu Thr Asp Tyr Ile Val Arg Glu Val Leu Leu Lys Tyr
675      680      685
Thr Ala Asp Pro Asp Gln His Lys Phe Leu Leu Phe Lys Lys Pro Val
690      695      700
Met Asp Leu Ile Ala Lys Ile Val Thr His Tyr Ala Val Ile His Ser
705      710      715
Ala Ala Asp Asn Gly Gly Val Cys Leu Ala Phe Pro Arg Asp Pro Pro
725      730      735
Phe Ile Val Glu Asn Asp Thr Ser Leu Arg Tyr Tyr Thr Leu Thr Asp
740      745      750
Thr Pro Gln Ser Ile Leu Asn Gly Asp Asn Val Ala Glu Asn Leu Lys
755      760      765
Ser Ala Thr Ser Val Ala Ser Ser Pro Ser Ser Ser Ser Arg Tyr Ser
770      775      780
Ser Glu Thr Pro Ile Arg Val Val Asn Leu Pro Val Pro Thr Gly Arg
785      790      795
Phe Leu Lys Met Asn Lys Asp Leu Glu Leu Phe Ile Asn Val Pro Leu
805      810      815
Ile Ser Ser Lys Glu Gln Lys Gln Gln Gln Gln Gln Thr Thr Ala Thr
820      825      830
Ala Pro Phe Ser Ser Glu Thr Ile Ser Lys Ser Phe Leu Asn Tyr Val
835      840      845
Pro Pro Lys Ser Leu Thr Arg Asn Val Thr Tyr Gly Gln Asn Ile Ala
850      855      860
Glu Asp Gly Phe Leu Gly Leu Lys Asn Lys Gly Glu Leu Val Ser Tyr
865      870      875
Phe Lys Val Val Lys Asn Thr Glu Arg Asp Asp Gly Ile Lys Asp Met
885      890      895
Glu Ile Gly Asp Ile Asn Asn His Gln Asp Asp Thr Gly Ser Leu Ser
900      905      910
Ser Ser Ser Ser Ser Phe Val Asp Gly Val Arg Thr Ser Phe Ser Val
915      920      925
Asp Gly Lys Ile Glu His Val Ser Ala Phe Leu Pro Gly Thr Thr Ser
930      935      940
Gln Pro Thr Asn Leu Pro Val His Ala Ser Lys Gln Val Lys Tyr Ser
945      950      955
Val Lys Glu Leu Gly Met Val Phe Phe Glu Pro Leu Leu Ser Ser Ala
965      970      975
Val Leu Tyr Glu Ala Ser Lys Thr Lys Ser Thr Gln His Leu Ser Pro
980      985      990
Met Arg Ile Tyr Lys Glu Cys Val Ser Pro Leu Ser Thr Gly Arg Ile
995      1000      1005

```

WO 01/38351

326

PCT/US00/28888

Asp Ile Phe Pro Ser Lys Val Gly Thr Val Ala Gly Thr Gly Phe Glu  
1010 1015 1020  
Phe Ile Trp Lys Val Leu Gln Tyr Asp Thr Gly Leu Pro Thr Thr Leu  
1025 1030 1035 1040  
Glu Arg Leu Ser Pro Lys Ile Pro Ser Val Pro Ile Ser Gly Glu Asp  
1045 1050 1055  
Ser Lys Met Glu Val Ile Ala Glu Ser Gly Lys Gly Val Gln Asn Ile  
1060 1065 1070  
Ile Ala Ile Ala Ala Asp Gln Leu Arg Gly Ser Asn Asn Ile Val Gly  
1075 1080 1085  
Gly Gly Thr Arg Arg Ala Ile Gln Gln Gln Gln Gln Gln Gln Gln  
1090 1095 1100  
Glu Thr Gln Ala Val Val Pro Val Asn Val Pro Ala Arg Phe Glu Pro  
1105 1110 1115 1120  
Thr Phe Thr Glu Ile Glu Leu Phe Leu Gln Asn Lys Phe Arg Asn Val  
1125 1130 1135  
Ile Ala Thr Ile Ile Ser Arg Met Met Met Leu Val Ser Asn Glu Glu  
1140 1145 1150  
Met Lys Ile Ile Lys Glu Val Cys Glu His Val Ser His Ile Met Val  
1155 1160 1165  
Asp Gly Val Asp Pro Arg Lys Ala Ile Glu Glu Ile Arg Ile Thr Ala  
1170 1175 1180  
Glu Gln Asn Gly Ile Thr Ile Asp Thr Gly Asn Glu Gly Tyr Gly Ser  
1185 1190 1195 1200  
Leu Arg Tyr Ala Ser Ser Gly Arg Leu Phe Ile Asn Asp Glu Ala Ser  
1205 1210 1215  
Glu Glu Ala Ala Ala Ile Gly Gly Gly Ala Leu Gly Thr Gly  
1220 1225 1230  
Arg Arg Val Pro Val Glu Leu Arg Ser Ile Leu Asp Lys Leu Asn Thr  
1235 1240 1245  
Ile Gly Ser Thr Thr Gln Gln Gln Gln Gln Gln Arg Gln Gln Arg  
1250 1255 1260  
Gln Ala Asn Asn Asn Asn Thr Val Pro Glu Asp Ile Lys Val His Asn  
1265 1270 1275 1280  
Glu Gln Met Gln Lys Ile Arg Asp Ser Ser Leu Phe Thr Ser Lys Leu  
1285 1290 1295  
Leu Asn Tyr Ile Arg Asp Asp Gly Arg Lys Asp Arg Ile Lys Thr Asn  
1300 1305 1310  
Ile Ser Glu Thr Leu Lys Lys Tyr Ser Arg Ile Pro Ser Tyr Phe Ile  
1315 1320 1325  
Ala Ser Lys Ala Gln Lys Pro Ile Pro Trp Lys His Thr Lys Asp Asn  
1330 1335 1340  
Ile Asn Leu Asn Lys Ile Pro Glu Asp Leu Asn Phe Ser Pro Ala Gln  
1345 1350 1355 1360  
Asn Leu Phe Val Pro Val Asn Pro Arg His Ile Leu Thr Asp Met Gln  
1365 1370 1375  
Trp Leu Asn Cys Ile Ser Ile Ile Glu Thr Ala Thr Arg Asp Ser Ala  
1380 1385 1390  
Ile Val Met Gln Ser Phe Gln Glu Gln Ala Asp Lys Thr Thr Thr Gln  
1395 1400 1405  
Leu Glu Glu Leu Leu Ser Gln Trp Asn Asn Ile Val Ser Gln Val Thr  
1410 1415 1420  
Asp Glu Lys Ser Pro Ala Tyr Val Ser Ser Val Lys Leu Glu Trp Leu  
1425 1430 1435 1440  
Asn Asn Glu Ala Ser Arg Ile Ala Ala Ile Arg Glu Asn Ser Glu Lys  
1445 1450 1455  
Ser Lys Ile Val Met Gly Val Gln Gly Lys Ile Val Asn Ile Asp Glu  
1460 1465 1470  
Leu Gly Ile Val Ala Val Ala Arg Ser Ile Val Asp Val Asp Phe Tyr  
1475 1480 1485  
Ile Lys Met Pro Asn Val Trp Ala Ser Arg Asp Trp Lys Asn Leu Ile

```

      1490              1495              1500
Tyr Tyr Ala Val Asn Ile Ala Ala Thr Pro Leu Ile Asn Asn Ile Ser
1505              1510              1515              1520
Arg Gly Ile Met Ala Ala Ser Gln Thr Ser Val Leu Tyr Asp Ser Ser
      1525              1530              1535
Leu Ala Leu Ile Ala Ala Glu Gln Ala Thr Arg Asn Ile Thr Met
      1540              1545              1550

```

<210> 236  
 <211> 669  
 <212> DNA  
 <213> SHRIMP

```

<400> 236
atggaatgga taaaccaacg gacaagtaga gaggatcttt tcaacacgta tacaggggaat 60
gcagtaatac gttcggcagc taagcaagca ctggctattg aaaaacacgc agcagaaaga 120
agaggagaaa aggcattggac gacttcagca gcagcagcag cttcttctaa ttttaataat 180
gtacaacaag attatactga tgatgatatt acacaagtgt ctattgcaaa cagtgttttg 240
aataaccctt ttttaagag atatgcaaaa cttatagata atttagcaat atcttcttta 300
cctcctgata tagaggatga tgtcattata cacactagag atgcctccaa ctctacagtc 360
agagtagatg gagccaatat ctatttcgcc ataattgacg gtgatttatg tgtataccct 420
aaacaatata tatctgataa agtgctgtgt ggttctctca accgggaaaa ggcactgttc 480
tataatagct ccaagaataa gtggacgtat ggatgtaacc taaactttga tatcgttgac 540
gctgccatca tgaacacccc cgactacaag gaagagacta catctacaaa acatatacgt 600
aaaatatttg gtatcggagc atcggaaaaa ctgaacatta cccactattt aaactacttt 660
atccaataa

```

<210> 237  
 <211> 218  
 <212> PRT  
 <213> SHRIMP

```

<400> 237
Met Glu Trp Ile Asn Gln Arg Thr Ser Arg Glu Asp Leu Phe Asn Thr
  1              5              10              15
Tyr Thr Gly Asn Ala Val Ile Arg Ser Ala Ala Lys Gln Ala Ile Glu
      20              25              30
Lys His Ala Ala Glu Arg Arg Gly Glu Lys Ala Trp Thr Thr Ser Ala
      35              40              45
Ala Ala Ala Ala Ser Ser Asn Phe Asn Asn Val Gln Gln Asp Tyr Thr
      50              55              60
Asp Asp Asp Ile Thr Gln Val Ser Ile Ala Asn Ser Val Leu Asn Asn
      65              70              75              80
Pro Phe Leu Lys Arg Tyr Ala Lys Leu Ile Asp Asn Leu Ala Ile Ser
      85              90              95
Ser Leu Pro Pro Asp Ile Glu Asp Asp Val Ile Ile His Thr Arg Asp
      100             105             110
Asn Ser Thr Val Arg Val Asp Gly Ala Asn Ile Tyr Phe Ala Ile Ile
      115             120             125
Asp Gly Asp Leu Cys Val Tyr Pro Lys Gln Tyr Ile Ser Asp Lys Val
      130             135             140
Leu Cys Gly Ser Leu Asn Arg Glu Lys Ala Leu Phe Tyr Asn Ser Ser
      145             150             155             160
Lys Asn Lys Trp Thr Tyr Gly Cys Asn Leu Asn Phe Asp Ile Val Asp
      165             170             175
Ala Ala Ile Met Lys His Pro Asp Tyr Lys Glu Glu Thr Thr Ser Thr
      180             185             190
Lys His Ile Arg Lys Ile Leu Gly Ile Gly Ala Ser Glu Lys Leu Asn
      195             200             205
Ile Thr His Tyr Leu Asn Tyr Phe Ile Gln

```

atagagcata	cagggcagat	ggaagaaaat	aggcctgcta	ccagaaaaag	gagacctgga	60
atgagaagaag	aggaagaaac	tggtagtagt	aatgttccat	attatgccaa	ctttggcgat	120
gacgccacgt	actccatgta	cactggagaa	ggaaaaaggg	gtaaatttgt	attagagcca	180
cctaaagaaa	gaagtgtaca	aaggggtgcaa	aaaccaccta	aagaaaagga	ggaaagggaa	240
caacgcttcta	atgttcggac	acggagacct	ggctcaagagt	ttgaacagaa	agtgctacaa	300
gatcgatcgc	gagaacggtc	agaaaaactt	gggcaaaatt	tggcagagaa	aggattgcaa	360
gaacggcaaa	agaaatatac	tccaaaggta	gcacaaacaa	tgacaaaaaa	aataatcagg	420
tttcgtgaag	gaggaagaaa	atccaaggcg	ccgcaacagc	agacatctga	caaaggtgca	480
gcaaccaatg	ttcttgaag	ggaagaaatt	gagatggctg	cagaaagaga	acaaccagta	540
gaaattacag	gagatactat	attaggctgg	ctaggagcag	aagatgcaga	agatatggga	600
gaggatgaat	taactataca	acattcatct	attggctgtat	caacaaccgt	tcaacaaatc	660
gttgtcagtt	ctcctatacc	gccaaagccc	actaggcccg	ctcctgatat	ttccatacaa	720
gaagatatag	tggggaaaaa	tattagccag	ttaccaccat	taccacttga	tgattatgag	780
gacgaagaag	acgaacattt	gtacgaagaa	gtgaatgatt	tcttagtgct	accaccaaca	840
gcagcagcag	cagcttccac	aaagacctcc	aggccttaata	ttcctcctcc	acctcctcct	900
gttgtgtctg	ttgcagacga	aaacctgaag	aacttggcct	caattgcagc	cttggaaaaag	960
gaggccgagg	aacaaagagc	ggccgcagtt	gaaagggaaa	gagaagtaga	ggaacaaaga	1020
gcggccgctg	ctgctgctgc	tgctgtctgt	gccgcccaac	gggaagcaga	cgaaaaaagg	1080
gaaagagaaag	cagaggaaca	aagagcggcc	gctgtctgct	ctgctgcgcg	tgctgcccaa	1140
cgggaagcag	acgaaaaaaa	ggaaagagaa	gtagaggaaac	aaagagcggc	cgctgtgctg	1200
ctgctgccc	ctgctgccca	acgggaagca	gacgaaaaaa	gggaagagaga	agtagaggaa	1260
caaaagagcg	ccgctgctgc	tgaaaaggaa	atacttgccc	aacaaacttca	agaaatgaaa	1320
gaacaaatgc	gcataaagga	agaggagagg	cggaaagaac	tagcagataa	ggaggaagaa	1380
aaacgtcgag	aactagcagc	caaggaggaa	gaaaagcgtc	aagaatatatt	agctaaagaa	1440
gagcaaatgt	aaaaattgaa	tttccagttg	ggtagcagaa	tcacgtccaa	aagagcactc	1500
gaacaaatgt	tagaagaaga	gaaggcctca	gcctcacggg	cccgctccag	tgcacagtta	1560
gcgatccaag	caatagaata	tgaggatgaa	cttcttcagg	cagtcgaacc	tcaaggacag	1620
ttagttccta	tggtacgga	tttgtacgga	aaaatgtacg	atctcaataa	gaaattagaa	1680
gtacagaata	atacattaac	ttctgcattt	gaagacgtga	acaaaaacaa	cgaacagAAC	1740
caatttggtg	ctcaatccct	tgaaaaaatc	gctaaagcca	ttgaaaaatt	aactagtcaa	1800
aaacatcttc	ctgtggatga	ctctgctttt	atgcagagaa	taataacaga	gagggatttt	1860
tcctttaaaga	atctgggaaa	tgtttacaaa	agagttctcg	gggtctatit	tacattaaaa	1920
agggaccttt	ttaaatcgaa	ggcatttaatt	acagataaag	aatcaaggga	tctggagggtg	1980
cgtctaacag	atgtatcgac	agatctcagg	gctaattgatc	tcaatacaat	actggaaagg	2040
ttggatgtat	ccgttaaacat	acgctctggt	ggaacattat	acactaaatt	tacagaggca	2100
gacacggcat	tagcagatca	agttccttcg	aggattgaaa	taagtaacag	atacaagtct	2160
gccttattgc	cattttcatc	tcagggtttg	gatactaatt	ttactaatag	ttccgacaag	2220
tacaatgaaa	tagtgaacca	actaagcagt	ataaatgagg	ctatgaatat	tttgaagaaa	2280
aatattgtcc	caacattgaa	ccaaatcaaa	attgatgtca	ccaatctatt	aacagtttca	2340
agctctcgtc	aatattgctat	tgaagaaaag	gtgtattctg	atgtgtcccg	aatggattct	2400
gaaataagaa	aatctctcgc	tataatgaac	agtaaaaatat	ccccttattt	taaggcgat	2460
tgacgggatg	aaagacaacg	ctctattgct	gacagtatit	cctctcagat	aaaatacaac	2520
gataaaaatta	aagagagtg	tgctacacta	cacgatataca	atacaacctc	aagaatacgt	2580
agtaatcccc	tctgcacaa	atcctcagtt	ttatcatctc	cagacttttt	aaatgctgtt	2640
aacgacttta	gaaattttct	cgatatccaa	ggaggttctc	aaatttactta	tgatgtcctt	2700
tcaggccaaa	atatttgctat	cgcttaccag	gcatacaaaa	ccactgaaaa	ggttacagaa	2760
ttgtgcctcg	aattatccat	aattttagac	gtgatccata	aaaatgcttt	gagttttaa	2820
ttgcctgcaa	ttacttacc	cgcaggggaa	acatctatgg	aagaaagtgg	ttcattggct	2880
gttgacatta	gacaagaaat	tggtaaagaa	atatcagatt	ctagcgcgtga	acttagctcg	2940
acactgtcag	aggcgtttaca	aatttttcag	caacaacagc	aacaacaaca	gcaacaattc	3000
caacagcagc	tggttacaaca	acaacaggac	caacaaatc	aacaacaatt	attacaaca	3060
caaatagaag	aacaacaacg	gqttcagaaa	caacagcaac	aqcaqcaaaq	qaccaacaa	3120

caacaggaac	aacagcaaag	ggaacaacag	cagcaacagc	agcaaaggga	acaacaacaa	3180
caaagggaac	aacaacagca	gcaacagcag	caacagagt	accagtttcg	acaacaatta	3240
ttgcaacaac	agcaacaatt	tcagcaatta	ctacaacaac	aagggaagaag	aaggaggggt	3300
gacgatggtg	atgaagaaag	agaggaaaga	gaagaagggg	ctgaaaagga	tgattgtgtg	3360
cgtaaggttg	cagaatcagt	agcgacaaaa	tatactgctg	acttgactac	cttattccaa	3420
cgagaagaaa	ataacttcca	atctaaaaata	gcatcagcaa	aattgggaac	ccttgtcttt	3480
gccacccctc	cttcacctat	catgaacttg	acaaaattga	gagaggaata	ttccacattc	3540
acaaccagtg	gtttttcaaa	actaacagct	gaaaataata	gtattatgcg	tattttcccc	3600
gagaggattg	tagaagtatg	caagagtaag	aatctcaatt	taatgggaaa	atacttgtac	3660
attataacta	ccgcacaaaac	agaaatggaa	gatcgagtga	agaatatatt	gtctggtatt	3720
ttcaatcaaa	ttgaagagtt	ttcaacaacat	gtaaaaacaac	aacaacaaca	acaagctgct	3780
tcagcttctt	ctactaatcc	tcctcctcct	tctactcctt	ctactactcc	tcctgttaca	3840
agcatgcaag	tttgtgagtt	ggatgatcaa	cgtaccctag	aaaaggctgc	catagtagag	3900
gcaattactc	tggccaatgc	tgtacttcaa	actacaaaat	ccgcttcagc	tccttccacg	3960
gcggctgagc	gagaaattgc	tctaaagcta	gagaatggga	aaacatctat	ccgtatggaa	4020
aaagtggatc	taagttcagg	agctactggt	gtttccgacc	aacaaaaatg	gatcgacgaa	4080
agtacttcca	aacaagaatt	ggaagatttc	attgcagaag	aaaactttgt	agaaactgca	4140
cataatgaaa	tgatatttgg	attaattttg	gatgccaaaga	agaacgatcc	gaccctgtat	4200
gccaatctta	ggctcgtgaa	acctcatgga	ataaatgtgc	agtctttccc	atattacgta	4260
ctccgcacat	ggctaggaga	aacagatata	ttggatgaag	atactgtaca	tcctgaatat	4320
ttccgcgaat	acattgatcg	caattggaag	gtggaagaac	atgagcgtga	agatacata	4380
aaggcactgg	gtgtttcttt	atcagatacg	ttggcacaca	tcaaggacta	ctattctccc	4440
agtgctcaaaa	atgatgcac	aaaatcagta	ccatttgcgt	tgaacactct	attgtacaac	4500
atatttgcta	tcgacggagg	aatgatttct	agcctttcaa	gaacagcctt	tatttaccgt	4560
aaatttttaa	ggcaatctat	gacagataaa	gaggttgctc	aaggacctgt	tcggtctcaa	4620
ctgtgtgaag	cgacaatagc	gtctcttttc	acggcatgta	gtaaccttct	tcgatcgtct	4680
cccttagccg	ataaagtaga	accacgcctt	caagaaaaat	tagcggccgc	tgctgctgta	4740
gacacgtcaa	ccggagacat	gtttcgtata	cgagtttgct	atctcatgta	caattttata	4800
gtggcttatg	tgaacctatg	caataatcgt	ataaactata	cgtaaagtgt	attgagagcg	4860
tcgggtctgg	caaaacaaaa	agtcgtggcc	ggtaaaacaa	ctaaagggca	tacatcttct	4920
tcccaccggg	ttggatctta	tgatgtcaca	tatgattttt	cagtattgta	caagattctt	4980
caactacaaa	aacagaatat	ctctttactt	ttagagaagg	gtttcaatgc	atgggaatcg	5040
tgtgttgccg	ccatggcagc	cttactgcc	gaccttctc	tctccatctc	agatgccgac	5100
caatccatac	tcttcccact	agaaggggga	gagattgtta	ttgaaaaaca	tgaaaacgat	5160
gcagaaaaaa	atgtcgatat	ggttcaagaa	ttgtggaagg	aaactgcact	cacactcatg	5220
gcaaaggaac	taaattcata	ctacaactgg	ttcatataca	gtaaggatac	cgatatggaa	5280
aagttggcga	gagtggtcag	gatgattatc	ggtatagtta	aggctgttct	cagattaaca	5340
aataaagccg	agagtttagt	agacactaat	gctctttctg	atatatttcaa	gctacctgtg	5400
attccaattg	atgataccaa	aactctggca	ataaatattg	tggtgttcac	tttgaataat	5460
gtaattaaac	catggatggt	ttcgttcaag	caaatgttta	gacagaaaga	cgggggagtt	5520
tcttcagcct	atttctcggt	tcaaaatata	caacaacaaa	aacatcaaca	aactgcttcc	5580
attttagatg	catgggcttg	cgcgcccggt	aagttaacaa	aggcagcaca	cgtgtttatt	5640
tcaggatatg	aaaatcatat	caaattaaaa	aaagatgata	ttttatgggg	tgcttcaatg	5700
aaattccctg	ccgacggacg	aggcactgtg	gtcgaaggat	gggccaaca	atacaataac	5760
gaaagcgtat	tggagagatt	taccgacttt	tcaattgaag	taaacgcacc	tgctctgga	5820
ctcttaatac	cgccagatcc	tttactttct	tccatgttcg	gtaaaggaaa	tggtggaagc	5880
agcagcagca	gcagcaagga	taatacaatt	attggaaagg	gaggattaat	tttaaaccgc	5940
caagttgttg	gacaagaaca	agcaccacca	ataaacacgt	cgtcggatac	taaaaaata	6000
agacgtgatg	caaataattga	accaataata	gggacgcctt	atagtgtaat	taaggcaagt	6060
aaaggagtat	caatttcagt	actggatgat	ttcaatgagg	atagtccaga	agatttcgcc	6120
cttaaaactt	ccatcatcaa	tgatgccata	cgagaaatag	ggcaacgcac	gacttatata	6180
agacctatat	ttgatcatca	aacacagaaa	aatatacact	acagttcacc	taaaattatt	6240
ctcgaaggat	catgatttaa	gaatggacaa	cgttcaggac	aatcttgggc	tcctcttcca	6300
tcctccttga	ctctggcctc	cgatttgaat	ctaccttctc	tggagctttt	atatagagaa	6360
cttgccacaa	aacaagtaga	gaagggaaga	gaagaaaaga	gcgaaaggga	agaagataaa	6420
ggacaaaaac	ttaatgaaaa	attatcattt	gtcgtgaata	aagctatcgg	aactatccaa	6480
caacaacatc	aatattctga	aaggggagga	ggaatgaaga	ggtatcagca	acactctgct	6540
gatcaagcta	gtaatggtgg	catagatgat	atagaactta	tgaatagtaa	agatgctact	6600
tccatgagaa	aggcaaaact	ggcattagcc	gttactaata	aaattgcagc	agcagcagca	6660
agggatgggg	aaaattcatc	agctaaaccg	tcaaaccttg	gcaatagatt	ggatgaagca	6720
ataaaccttg	gagcactttt	attacgtaga	ggaggaggag	taagaggagg	acaaacaccc	6780

WO 01/38351

330

PCT/US00/28888

<210> 239  
<211> 2293  
<212> PRT  
<213> SHRIMP

<400> 239

Met	Ser	Asp	Thr	Gly	Gln	Met	Glu	Glu	Asn	Arg	Pro	Ala	Thr	Gln	Lys
1				5					10					15	
Arg	Arg	Pro	Gly	Asp	Glu	Glu	Glu	Glu	Glu	Thr	Gly	Ser	Ser	Asn	Val
			20					25					30		
Pro	Tyr	Tyr	Ala	Asn	Phe	Gly	Asp	Asp	Ala	Thr	Tyr	Ser	Met	Tyr	Thr
		35					40					45			
Gly	Glu	Gly	Lys	Arg	Gly	Lys	Phe	Val	Leu	Glu	Pro	Pro	Lys	Glu	Arg
	50					55					60				
Ser	Val	Gln	Arg	Val	Gln	Lys	Pro	Pro	Lys	Glu	Lys	Glu	Glu	Arg	Glu
65					70					75				80	
Gln	Arg	Ser	Asn	Val	Arg	Thr	Arg	Arg	Pro	Gly	Gln	Glu	Phe	Glu	Gln
			85						90					95	
Lys	Val	Leu	Gln	Asp	Arg	Ser	Arg	Glu	Arg	Ser	Glu	Lys	Leu	Gly	Gln
			100					105					110		
Asn	Leu	Ala	Glu	Lys	Gly	Leu	Gln	Glu	Arg	Gln	Lys	Lys	Tyr	Thr	Pro
	115						120					125			
Lys	Val	Ala	Gln	Thr	Met	Thr	Lys	Lys	Ile	Ile	Arg	Phe	Arg	Glu	Gly
	130				135						140				
Gly	Arg	Lys	Phe	Lys	Ala	Pro	Gln	Gln	Gln	Thr	Ser	Asp	Lys	Gly	Ala
145					150					155					160
Ala	Thr	Asn	Val	Leu	Glu	Arg	Glu	Glu	Ile	Glu	Met	Ala	Ala	Glu	Arg
				165					170					175	
Glu	Gln	Pro	Val	Glu	Ile	Thr	Gly	Asp	Thr	Ile	Leu	Gly	Gly	Leu	Gly
			180				185					190			
Glu	Glu	Asp	Asp	Glu	Asp	Met	Gly	Glu	Asp	Glu	Leu	Thr	Ile	Gln	His
	195						200					205			
Ser	Ser	Met	Ala	Val	Ser	Gln	Pro	Val	Gln	Gln	Ile	Val	Val	Ser	Ser
	210					215					220				
Pro	Ile	Pro	Pro	Lys	Pro	Trp	Ala	Pro	Asp	Ile	Pro	Ile	Gln	Glu	Asp
225					230				235						240
Ile	Val	Gly	Lys	Asn	Ile	Ser	Gln	Leu	Pro	Pro	Leu	Pro	Leu	Asp	Asp
				245					250					255	
Tyr	Glu	Asp	Glu	Glu	Asp	Glu	His	Leu	Tyr	Glu	Glu	Val	Asn	Asp	Phe
		260						265					270		
Leu	Val	Ala	Pro	Pro	Thr	Ala	Ala	Ala	Ala	Ala	Ser	Trp	Pro	Arg	Pro
		275					280					285			
Asn	Ile	Pro	Pro	Pro	Pro	Pro	Pro	Val	Val	Ala	Val	Ala	Asp	Glu	Thr
	290					295					300				
Leu	Lys	Asn	Leu	Ala	Ser	Ile	Ala	Ala	Leu	Glu	Lys	Glu	Ala	Glu	Glu
305					310					315					320
Gln	Arg	Ala	Ala	Ala	Val	Glu	Arg	Glu	Arg	Glu	Val	Glu	Glu	Gln	Arg
				325					330					335	
Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gln	Arg	Glu	Ala
			340					345					350		
Asp	Glu	Lys	Arg	Glu	Arg	Glu	Ala	Glu	Glu	Gln	Arg	Ala	Ala	Ala	Ala
		355					360					365			
Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gln	Arg	Glu	Ala	Asp	Glu	Lys	Lys	Glu
	370					375					380				
Arg	Glu	Val	Glu	Glu	Gln	Arg	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala
385					390					395					400
Ala	Ala	Gln	Arg	Glu	Ala	Asp	Glu	Lys	Arg	Glu	Arg	Glu	Val	Glu	Glu

405 410 415  
 Gln Arg Ala Ala Ala Ala Glu Arg Glu Ile Leu Ala Gln Gln Leu  
 420 425 430  
 Gln Glu Met Lys Glu Gln Met Arg Ile Lys Glu Glu Glu Arg Arg Lys  
 435 440 445  
 Glu Leu Ala Asp Lys Glu Glu Lys Arg Arg Glu Leu Ala Ala Lys  
 450 455 460  
 Glu Glu Glu Lys Arg Gln Glu Ile Leu Ala Lys Glu Glu Gln Leu Glu  
 465 470 475 480  
 Lys Leu Asn Phe Gln Leu Gly Thr Glu Ile Thr Ser Lys Arg Ala Leu  
 485 490 495  
 Glu Gln Met Leu Glu Glu Glu Lys Ala Ser Arg Ser Arg Ser Ala Ala  
 500 505 510  
 Ser Ala Gln Leu Ala Ile Gln Ala Ile Glu Tyr Glu Asp Glu Leu Pro  
 515 520 525  
 Gln Ala Val Glu Pro Gln Gly Gln Leu Val Pro Met Asp Thr Asp Leu  
 530 535 540  
 Tyr Gly Lys Met Tyr Asp Leu Asn Lys Lys Leu Glu Val Gln Asn Asn  
 545 550 555 560  
 Thr Leu Thr Ser Ala Phe Glu Asp Val Asn Lys Thr Asn Glu Gln Asn  
 565 570 575  
 Gln Leu Val Ala Gln Ser Leu Glu Lys Ser Ala Lys Ala Ile Glu Lys  
 580 585 590  
 Leu Thr Ser Gln Lys His Leu Pro Val Asp Asp Pro Ala Phe Met Gln  
 595 600 605  
 Arg Ile Ile Thr Glu Arg Asp Phe Ser Leu Lys Asn Leu Gly Asn Val  
 610 615 620  
 Tyr Lys Arg Val Leu Gly Val Tyr Phe Thr Leu Lys Arg Asp Leu Phe  
 625 630 635 640  
 Lys Ser Lys Ala Leu Ile Thr Asp Lys Glu Ser Arg Asp Leu Glu Val  
 645 650 655  
 Arg Leu Thr Asp Val Ser Thr Asp Leu Arg Ala Asn Asp Leu Asn Thr  
 660 665 670  
 Ile Arg Leu Asp Val Ser Val Asn Ile Arg Ser Gly Gly Thr Leu Tyr  
 675 680 685  
 Thr Lys Phe Thr Glu Ala Asp Thr Ala Asp Gln Val Pro Ser Arg Ile  
 690 695 700  
 Glu Ile Ser Asn Arg Ser Arg Ser Ala Leu Leu Pro Phe Ser Ser Ala  
 705 710 715 720  
 Gly Leu Asp Thr Asn Phe Thr Asn Ser Ser Asp Lys Tyr Asn Glu Ile  
 725 730 735  
 Val Asn Gln Leu Ser Ser Ile Asn Glu Ala Met Asn Ile Leu Lys Glu  
 740 745 750  
 Asn Ile Val Pro Thr Leu Asn Gln Ile Lys Ile Asp Val Thr Asn Leu  
 755 760 765  
 Leu Thr Val Ser Ser Ser Arg Gln Tyr Ala Ile Glu Glu Arg Val Tyr  
 770 775 780  
 Ser Asp Val Ser Arg Met Asp Ser Glu Ile Arg Lys Phe Leu Ala Ile  
 785 790 795 800  
 Met Asn Ser Lys Ile Ser Pro Tyr Phe Lys Gly Asp Trp Thr Asp Glu  
 805 810 815  
 Arg Gln Arg Ser Ile Ala Asp Ser Ile Ser Ser Gln Ile Lys Ser Asn  
 820 825 830  
 Asp Lys Ile Lys Glu Ser Val Ala Thr Leu His Asp Ile Asn Thr Thr  
 835 840 845  
 Ser Arg Ile Arg Ser Asn Pro Leu Leu His Lys Ser Ser Val Leu Ser  
 850 855 860  
 Ser Pro Asp Phe Leu Asn Ala Val Asn Asp Phe Arg Asn Phe Leu Asp  
 865 870 875 880  
 Ile Gln Gly Gly Ser Gln Phe Thr Tyr Asp Val Leu Ser Gly Gln Asn  
 885 890 895

Ile	Asp	Asp	Leu	Ser	Leu	Ala	Ser	Lys	Thr	Thr	Glu	Lys	Val	Thr	Glu
Leu	Cys	Leu	Glu	Leu	Ser	Ile	Ile	Leu	Asp	Val	Ile	His	Lys	Asn	Ala
Leu	Ser	Leu	Asn	Leu	Pro	Ala	Ile	Thr	Tyr	Pro	Ala	Gly	Glu	Thr	Ser
Met	Glu	Glu	Ser	Gly	Ser	Leu	Ala	Val	Asp	Ile	Arg	Gln	Glu	Ile	Gly
945	Lys	Asn	Ile	Ser	Asp	Ser	Ala	Glu	Leu	Ser	Arg	Thr	Leu	Ser	Glu
Ala	Leu	Gln	Ile	Phe	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Phe
Gln	Gln	Gln	Leu	Leu	Gln	Gln	Gln	Gln	Asp	Gln	Gln	Asn	Gln	Gln	Gln
Leu	Leu	Gln	Gln	Gln	Ile	Glu	Glu	Gln	Gln	Arg	Val	Gln	Glu	Gln	Gln
1010	Gln	Gln	Gln	Arg	Asp	Gln	Gln	Gln	Gln	Glu	Gln	Gln	Gln	Arg	Glu
1025	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Arg	Glu	Gln	Gln	Gln	Gln	Arg	Glu
Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Arg	Glu	Gln	Gln	Gln	Gln	Arg	Glu
Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Ser	Asp	Gln	Phe	Arg	Gln	Gln
Leu	Gln	Gln	Gln	Gln	Gln	Gln	Phe	Gln	Gln	Leu	Leu	Gln	Gln	Gln	Gly
1075	Arg	Arg	Gly	Gly	Asp	Asp	Gly	Asp	Glu	Glu	Arg	Glu	Glu	Arg	Glu
1090	Gly	Ala	Glu	Lys	Asp	Asp	Cys	Val	Arg	Lys	Val	Ala	Glu	Ser	Val
1105	Thr	Lys	Tyr	Thr	Ala	Asp	Leu	Thr	Thr	Leu	Phe	Gln	Arg	Glu	Glu
1115	Asn	Phe	Gln	Ser	Lys	Ile	Ala	Ser	Ala	Lys	Leu	Gly	Thr	Leu	Val
1125	Ala	Thr	Pro	Pro	Ser	Pro	Ile	Met	Asn	Leu	Thr	Lys	Leu	Arg	Glu
1135	Tyr	Ser	Thr	Phe	Thr	Thr	Gln	Cys	Phe	Ser	Lys	Leu	Thr	Ala	Glu
1145	Asn	Ser	Ile	Met	Arg	Ile	Phe	Pro	Glu	Arg	Ile	Val	Glu	Val	Cys
1155	Ser	Lys	Asn	Leu	Asn	Leu	Met	Gly	Lys	Tyr	Leu	Tyr	Ile	Ile	Thr
1165	Ala	Gln	Thr	Glu	Met	Glu	Asp	Arg	Val	Lys	Asn	Ile	Leu	Ser	Gly
1175	Phe	Asn	Gln	Ile	Glu	Glu	Phe	Ser	Asn	Asn	Val	Lys	Gln	Gln	Gln
1185	Gln	Gln	Ala	Ala	Ser	Ala	Ser	Ser	Thr	Asn	Pro	Pro	Pro	Pro	Ser
1190	Pro	Ser	Thr	Thr	Pro	Pro	Val	Thr	Ser	Met	Gln	Val	Cys	Glu	Leu
1200	Asp	Gln	Arg	Thr	Leu	Glu	Lys	Ala	Ala	Ile	Val	Glu	Ala	Ile	Thr
1210	Ala	Asn	Ala	Val	Leu	Gln	Thr	Thr	Lys	Ser	Ala	Ser	Ala	Pro	Ser
1220	Ala	Ala	Glu	Arg	Glu	Ile	Ala	Leu	Lys	Leu	Glu	Asn	Gly	Lys	Thr
1230	Ile	Arg	Met	Glu	Lys	Val	Asp	Leu	Ser	Ser	Gly	Ala	Thr	Gly	Val
1240	Asp	Gln	Gln	Lys	Trp	Ile	Asp	Glu	Ser	Thr	Ser	Lys	Gln	Glu	Leu
1250	Asp	Phe	Ile	Ala	Glu	Asn	Phe	Val	Glu	Thr	Ala	His	Asn	Glu	Met
1260	Asp	Ile	Gly	Leu	Ile	Leu	Asp	Ala	Lys	Lys	Asn	Asp	Pro	Thr	Arg
1270	Asp	Ile	Gly	Leu	Ile	Leu	Asp	Ala	Lys	Lys	Asn	Asp	Pro	Thr	Arg
1280	Asp	Ile	Gly	Leu	Ile	Leu	Asp	Ala	Lys	Lys	Asn	Asp	Pro	Thr	Arg
1290	Asp	Ile	Gly	Leu	Ile	Leu	Asp	Ala	Lys	Lys	Asn	Asp	Pro	Thr	Arg
1300	Asp	Ile	Gly	Leu	Ile	Leu	Asp	Ala	Lys	Lys	Asn	Asp	Pro	Thr	Arg
1310	Asp	Ile	Gly	Leu	Ile	Leu	Asp	Ala	Lys	Lys	Asn	Asp	Pro	Thr	Arg
1320	Asp	Ile	Gly	Leu	Ile	Leu	Asp	Ala	Lys	Lys	Asn	Asp	Pro	Thr	Arg
1330	Asp	Ile	Gly	Leu	Ile	Leu	Asp								



WO 01/38351

333

PCT/US00/28888

Ala	Asn	Leu	Arg	Leu	Val	Lys	Pro	His	Gly	Ile	Asn	Val	Gln	Ser	Phe
		1380						1385					1390		
Pro	Tyr	Tyr	Val	Leu	Arg	Thr	Trp	Leu	Gly	Glu	Thr	Asp	Ile	Leu	Asp
		1395						1400					1405		
Glu	Asp	Thr	Val	His	Pro	Glu	Tyr	Phe	Arg	Gln	Tyr	Ile	Asp	Arg	Asn
		1410						1415					1420		
Trp	Lys	Val	Glu	Glu	His	Glu	Arg	Glu	Asp	Thr	Leu	Lys	Ala	Leu	Gly
				1425		1430				1435					1440
Val	Ser	Leu	Ser	Asp	Thr	Leu	Ala	His	Ile	Lys	Asp	Tyr	Tyr	Ser	Pro
				1445					1450						1455
Ser	Val	Lys	Asn	Asp	Ala	Ser	Lys	Ser	Val	Pro	Phe	Ala	Leu	Asn	Thr
				1460					1465						1470
Leu	Leu	Tyr	Asn	Ile	Phe	Ala	Ile	Asp	Gly	Gly	Met	Ile	Ser	Ser	Leu
				1475					1480						1485
Ser	Arg	Thr	Ala	Phe	Ile	Tyr	Arg	Lys	Phe	Leu	Arg	Gln	Ser	Met	Thr
				1490					1495						1500
Asp	Lys	Glu	Val	Ala	Gln	Gly	Pro	Val	Arg	Ser	Gln	Leu	Cys	Glu	Ala
				1505					1510						1515
Thr	Ile	Ala	Ser	Leu	Phe	Thr	Ala	Cys	Ser	Asn	Leu	Leu	Arg	Ser	Ser
				1525					1530						1535
Pro	Leu	Ala	Asp	Lys	Val	Glu	Pro	Arg	Leu	Gln	Glu	Lys	Leu	Ala	Ala
				1540					1545						1550
Ala	Ala	Ala	Val	Asp	Thr	Ser	Thr	Gly	Asp	Met	Phe	Arg	Ile	Arg	Val
				1555					1560						1565
Cys	His	Leu	Met	Tyr	Asn	Phe	Ile	Val	Ala	Tyr	Val	Asn	Leu	Cys	Asn
				1570					1575						1580
Asn	Arg	Ile	Asn	Tyr	Thr	Leu	Asn	Val	Leu	Arg	Ala	Ser	Gly	Leu	Ala
				1585					1590						1595
Asn	Lys	Lys	Val	Val	Ala	Gly	Lys	Thr	Thr	Lys	Gly	His	Thr	Ser	Ser
				1605					1610						1615
Ser	His	Arg	Phe	Gly	Ser	Tyr	Asp	Val	Thr	Tyr	Asp	Phe	Ser	Val	Leu
				1620					1625						1630
Tyr	Lys	Ile	Leu	Gln	Leu	Gln	Lys	Gln	Asn	Ile	Ser	Leu	Leu	Leu	Glu
				1635					1640						1645
Lys	Gly	Phe	Asn	Ala	Trp	Glu	Ser	Cys	Val	Ala	Ala	Met	Ala	Ala	Phe
				1650					1655						1660
Thr	Ala	Asp	Pro	Ser	Leu	Ser	Ile	Ser	Asp	Ala	Asp	Gln	Ser	Ile	Leu
				1665					1670						1675
Phe	Pro	Leu	Glu	Gly	Gly	Glu	Ile	Val	Ile	Glu	Lys	His	Glu	Asn	Asp
				1685					1690						1695
Ala	Glu	Lys	Asn	Val	Asp	Met	Val	Gln	Glu	Leu	Trp	Lys	Glu	Thr	Ala
				1700					1705						1710
Leu	Thr	Leu	Met	Ala	Lys	Glu	Leu	Asn	Ser	Tyr	Tyr	Asn	Trp	Phe	Ile
				1715					1720						1725
Tyr	Ser	Lys	Asp	Thr	Asp	Met	Glu	Lys	Leu	Ala	Arg	Val	Cys	Arg	Met
				1730					1735						1740
Ile	Ile	Gly	Ile	Val	Lys	Ala	Val	Leu	Arg	Leu	Thr	Asn	Lys	Ala	Glu
				1745					1750						1755
Ser	Leu	Val	Asp	Thr	Asn	Ala	Leu	Ser	Asp	Ile	Phe	Lys	Leu	Pro	Val
				1765					1770						1775
Ile	Pro	Ile	Asp	Asp	Thr	Lys	Thr	Leu	Ala	Ile	Asn	Ile	Val	Val	Phe
				1780					1785						1790
Thr	Leu	Asn	Asn	Val	Ile	Lys	Pro	Trp	Met	Val	Ser	Phe	Lys	Gln	Met
				1795					1800						1805
Phe	Arg	Gln	Lys	Asp	Gly	Gly	Val	Ser	Ser	Ala	Tyr	Phe	Ser	Phe	Gln
				1810					1815						1820
Asn	Ile	Gln	Gln	Gln	Lys	His	Gln	Gln	Thr	Ala	Ser	Ile	Leu	Asp	Ala
				1825					1830						1835
Trp	Ala	Cys	Ala	Pro	Gly	Lys	Leu	Thr	Lys	Ala	Ala	His	Val	Phe	Ile
				1845					1850						1855
				1860					1865						1870

Ser Gly Tyr Glu Asn His Ile Lys Leu Lys Lys Asp Asp Leu Leu Trp  
 1875 1880 1885  
 Gly Ala Ser Met Lys Phe Pro Ala Asp Gly Arg Gly Thr Val Val Glu  
 1890 1895 1900  
 Gly Trp Ala Gln Gln Tyr Asn Asn Glu Ser Val Leu Glu Asp Phe Thr  
 1905 1910 1915 1920  
 Asp Phe Ser Ile Glu Val Asn Ala Pro Ala Ser Gly Leu Leu Ile Pro  
 1925 1930 1935  
 Pro Asp Pro Leu Leu Ser Ser Met Phe Gly Lys Gly Asn Gly Gly Ser  
 1940 1945 1950  
 Ser Ser Ser Ser Ser Lys Asp Asn Thr Ile Ile Gly Lys Gly Gly Leu  
 1955 1960 1965  
 Ile Leu Asn Arg Gln Val Val Gly Gln Glu Gln Ala Pro Pro Ile Asn  
 1970 1975 1980  
 Thr Ser Ser Asp Thr Lys Lys Ile Arg Arg Asp Ala Asn Ile Glu Pro  
 1985 1990 1995 2000  
 Ile Ile Gly Thr Pro Tyr Ser Val Ile Lys Ala Ser Lys Gly Val Ser  
 2005 2010 2015  
 Ile Ser Val Leu Asp Asp Phe Asn Glu Asp Ser Pro Glu Asp Phe Ala  
 2020 2025 2030  
 Leu Lys Thr Ser Ile Ile Asn Asp Ala Ile Arg Glu Ile Gly Gln Arg  
 2035 2040 2045  
 Met Thr Tyr Trp Ile Phe Asp His Gln Thr Gln Lys Asn Ile His Tyr  
 2050 2055 2060  
 Ser Ser Pro Lys Ile Ile Gly Ser Asp Leu Lys Asn Gly Gln Arg Ser  
 2065 2070 2075 2080  
 Gly Gln Ser Trp Ala Pro Ser Ser Ser Ser Leu Thr Leu Ala Ser Asp  
 2085 2090 2095  
 Trp Asn Leu Pro Ser Leu Glu Leu Leu Tyr Arg Glu Leu Ala Thr Lys  
 2100 2105 2110  
 Gln Val Glu Lys Glu Glu Glu Glu Lys Ser Glu Glu Asp Lys Gly Gln  
 2115 2120 2125  
 Lys Leu Asn Glu Lys Leu Ser Phe Val Val Asn Lys Ala Ile Gly Thr  
 2130 2135 2140  
 Ile Gln Gln Gln His Gln Tyr Ser Gly Gly Gly Met Lys Arg Tyr Gln  
 2145 2150 2155 2160  
 Gln His Ser Ala Asp Gln Asn Gly Gly Ile Asp Asp Ile Glu Leu Met  
 2165 2170 2175  
 Asn Ser Lys Asp Ala Thr Ser Met Arg Lys Ala Lys Leu Ala Val Thr  
 2180 2185 2190  
 Asn Lys Ile Ala Ala Ala Ala Ala Arg Asp Gly Glu Asn Ser Ser Ala  
 2195 2200 2205  
 Lys Pro Ser Asn Phe Gly Asn Arg Leu Asp Glu Ala Ile Asn Pro Gly  
 2210 2215 2220  
 Ala Leu Leu Leu Arg Arg Gly Gly Gly Val Arg Gly Gly Gln Thr Pro  
 2225 2230 2235 2240  
 Gln Ser Ser Met Leu Thr Met Phe Arg Pro Gly Gln Thr Gly Gly Asn  
 2245 2250 2255  
 Ser Ser Trp Trp Thr Thr Asn Thr Pro Leu Ile Gln Arg Thr Thr Ser  
 2260 2265 2270  
 Val Gly Asn Asn Leu Val Val Leu Val Pro Asn Leu Leu Asp Ser His  
 2275 2280 2285  
 Pro Pro Thr Phe Asn  
 2290

&lt;210&gt; 240

&lt;211&gt; 4311

&lt;212&gt; DNA

&lt;213&gt; SHRIMP

<400> 240

atggattttg	aaggaactac	cagttctacc	ccctcaaaaa	tgtcccagtt	gtattcatca	60
gtgaagaaag	ttgcagagca	ttcctttgcc	aatcttcatg	acaaggctac	tcttgcacat	120
aaggttatta	aggacctgga	aggggagagg	aagaaaatgt	ctaccccaaa	gtcctcttct	180
gatggacaaa	aactggacaa	ggctatgttg	gacgatatta	tcaacgagta	tcaggccgtt	240
aagagcactg	cagataattc	cattgaatcg	accatcaagg	aaattgaaaa	tgtacttgaa	300
agtgtgcgca	gaaccaagat	tgaaagtga	gccaagaaca	gtgtaacttc	cagcccagaa	360
aaagtgtttt	ctgtcgagga	tttagaaatc	tactccaagg	ggcgagtgtg	caaaggtctc	420
aagttaaaccg	ccaactgttc	aagaattgga	ggcaagtatg	cagtgtcaat	gagtatcaaa	480
aaacacaacg	tctcctcatt	tgagaacaac	aacaaccaag	ttttctctga	agaaccagg	540
gattgtttta	tgcttgaaac	aacctatcct	cttgttgggt	tcgaaacttc	tacagaagat	600
ggaaatacat	atgcagtttt	cttgactggg	gttgggctag	aaagatctct	acctaaatat	660
gtaccagttt	tcgacatgaa	tcgaggtatt	caaaccctaa	acatgactgg	tttgaggatg	720
gccaaagcttc	ctgttctgtg	catgtttgga	cgtacagaat	atgacaactt	ggaagatttt	780
tacatcactt	caattgagac	gcagtccttt	gacgaagagg	aaaatgatgc	cagaatgagg	840
tgtcacaccg	aaagatttga	gaggaagaag	cgcattgaatg	acgcaccagc	gattacacct	900
catgtggccg	tgtacgacta	cagtggagac	gggaagaac	aatgtctcta	tatgataacc	960
gagtatgaaa	acacggctag	ttggtgcaac	gcaaaccgtg	tggtcacatc	tgacagtggg	1020
ttttctaacc	aatgtgcaat	tagtgatatg	aatgacttgt	gctgttttgc	tgactgcatc	1080
gatgttactg	ttaataatga	agaacatgaa	gaacgttcta	tgaatatgtt	ggtcgaatct	1140
gacaggcgct	tttttgatgc	tagtccttcc	cccctcaaga	cggagaaga	tgagaaaaat	1200
tcactcgtcat	cgtcttcttc	tccaacagtt	cctctcctta	caccatacga	aggtaaacga	1260
gttgtggagg	gggaggaaga	agaggaagaa	attgatgaag	acgaaagtag	caagtatgaa	1320
ggttcagaag	atgctcttgt	tatgaagaaa	ttagccaagc	tttctactat	gaaacaaatg	1380
agaagggtta	agaatgaacc	tgcaactcaa	attacttctg	ggggtaacaa	tagtagcagt	1440
agtatcaata	acgaagatga	tggtgatgat	gacgatgccg	ttgacgctac	tgcatattgc	1500
ccccaaactg	aagctacagt	gaaaaatttc	ttcatggccc	caaacgacga	gagaactgaa	1560
aatatttttgt	atgaaactat	gcaaatttct	cttgctaaaa	tttgaataaa	tccatcatct	1620
atgagcagtt	accgtgtatt	caccaacaaa	ctccaagagt	gtttgaatac	catggacgat	1680
agtatccgct	gccgtccaac	catttggact	gaagaaagtc	aacaatttgc	taagggtttg	1740
ttgtttgatg	aggttgtcac	atcaatttgt	gcacatcaga	tggtcaaga	tatttgcaag	1800
tctgaaatat	ttggaggaat	gtttaacgcc	aactctacca	acattaaggg	taaatatgaa	1860
ggacaaaaga	agagtctgta	tgaaacaacg	cacatttctt	cctcgtgctt	caaaaccaac	1920
acggaatcta	atgtgaataa	tgcaactatt	gcgtgggtga	aatcgaaact	ccattctggc	1980
acagtcatac	ctaacgtatt	ctccttcaaa	atggcatcag	aaaagccctc	aaaaatgaag	2040
cgcaagcgta	cctctagtgc	ttcatcatct	aacgatgaac	accaagaacc	atcaacaaaa	2100
atgatgaaaa	atgatgaagg	ggaaaagggt	gcacaagaat	catcatctcc	ttcttcatcg	2160
tctacaccag	aaacaacaca	acaagctggt	catgacaagg	aaactatcaa	tttaattccc	2220
ctcagtttca	taaaaatgcc	acgcagtaat	gtcaatggct	cggtctcata	tttgtctgaa	2280
atattcgggtc	aacgtctttg	tggactgtct	gatgcttcca	gcacatttaa	gagaatgtgc	2340
aagacttttg	aagatcttga	aaatgaaatc	atgaggagct	cattcactag	actgactaga	2400
tatgagaggg	aggttaactcg	cttgtatgag	aaatgcaggt	ctcaagctgt	agatatggag	2460
gaaaatgaaa	tggatgtttt	gtctcaccac	ggggaattgt	ttgcccaggt	cttgaggagc	2520
ccgatcgctt	actttgaaga	agtactggag	aatattaaga	gttggagcct	agaaaacggt	2580
aacaccctta	agcgcaaaaa	caagtatgca	aaggtactgg	tgagcggtta	tgctattcgt	2640
aggacatatg	aagaatacca	tgcgtttagc	aagtttgtac	caatgttctt	gttcaatctg	2700
attaagagag	aattggaagg	agacaactat	acctatgacg	ttcacttttc	ttccacttgc	2760
ctgtggtacc	tgactgtaat	gaccaggaac	aggatttgcg	atgtgctcca	gtacatcaac	2820
aacaacaata	atgataacga	agaaaccgat	attgtggagg	aagaggagga	aggagaagga	2880
gaggaggata	aaatggaaga	aagtatggac	gtagaacaac	agaagcaagt	tcgcaaggga	2940
gggagaaaaga	agggtcaaaa	attcaacagt	attggggatc	aagtcattag	aaaattttgt	3000
aaaagtttgt	gtgaaaattc	gatggtagtt	tctattgcaa	ttaatagttt	gatctctgga	3060
ataagctgga	tgaacaagaa	aatccctccc	ggtttcttga	aggattctag	cacaatcaat	3120
acctttgatg	aggtctcaag	gtttgtgttt	agcgatgtca	aaatcaatag	gaaaatcaat	3180
ggaacagatg	ataaatatga	aactgttttt	ggagtcagta	cgcgtgtgga	ttcacatatt	3240
gtaggccctt	ttagtatacc	tgttgatttt	tcaagcgtag	gactagataa	ggcctcatgt	3300
ggcaaattgt	acgttaaacac	catagacgga	aagggcattt	tgacaatttc	acccaaatat	3360
gttccatttaa	acgatgagga	tgttgattct	actacaacag	acaagctaga	gaaggatatt	3420
tatgcatttgt	ctaagcatga	caccttttct	aatattaata	agaataaggt	tcttccattc	3480
tataatatatt	ctcctagctc	ttctctcact	gaaaagaaaa	agacaaaaat	caataggaag	3540
aagatctcat	ctggtatgag	caataataat	ggcatgtgtg	tacaaactcc	ttctagttca	3600

WO 01/38351

PCT/US00/28888

336

```

aattcagtct cttccgtctc gtctattgta gtccttcat cttctgttct ggctctatct 3660
tgctcccttt cttctacaaa gaaaaagagc atctggaacg agaacatgtt tttgacatct 3720
aggaacatgt ggaggtgtgg atttgttgta ccacccaaac tttgcagttt tattgttaac 3780
catagacacg ctgtaaaact tgtagctgaa actgcaccta aaacaaagtt gtgtaggaat 3840
attattgata ggaataggaa gattagattt aacggtctaa agaaggatg caagagtgtt 3900
agcgccctta cggcgagtc tacatatttg ctcaataaga atatgactgc aacttcacct 3960
agtgatttga acctatgtat ttatacttca tctttaaatg acccattgta tacttgtaaa 4020
ttgacccatg aagagtacca ggacggaaat gctttggatg actatggtgc agtttttgta 4080
aactatacat ttaagagtat caaatcctgc tcgtccaaag atgaaaccgc tgacgacaat 4140
gctgctgctg ctgatgatga tggttctact acttctacct cttcttcaac tgatactgat 4200
gctgctgcaa tccaagactt tatgcatgta atgattaaga aaattgatgc catgaaagac 4260
attagaggaa agtacaagaa atctttggca aagaaaacaa agaagcatta a 4311

```

<210> 241  
 <211> 1430  
 <212> PRT  
 <213> SHRIMP

<400> 241

```

Met Asp Phe Glu Gly Thr Thr Ser Ser Thr Pro Ser Lys Met Ser Gln
1      5      10      15
Lys Ser Val Lys Lys Val Ala Glu His Ser Phe Ala Asn Leu His Asp
20     25     30
Lys Ala Thr Leu Ala Ser Lys Val Ile Lys Asp Leu Glu Gly Glu Arg
35     40     45
Lys Lys Met Ser Thr Pro Lys Ser Ser Ser Asp Gly Gln Lys Leu Asp
50     55     60
Lys Ala Met Leu Asp Asp Ile Ile Asn Glu Tyr Gln Ala Val Lys Ser
65     70     75     80
Thr Ala Asp Asn Ser Ile Glu Ser Thr Ile Lys Glu Ile Glu Asn Val
85     90     95
Leu Glu Ser Val Arg Arg Thr Lys Ile Glu Ser Glu Ala Lys Asn Ser
100    105    110
Val Thr Ser Ser Pro Glu Lys Val Phe Ser Val Glu Asp Leu Glu Ile
115    120    125
Tyr Ser Lys Gly Arg Val Cys Lys Gly Leu Lys Leu Asn Ala Asn Cys
130    135    140
Ser Arg Ile Gly Gly Lys Tyr Ala Val Ser Met Ser Ile Lys Lys His
145    150    155    160
Asn Val Ser Ser Phe Glu Asn Asn Asn Asn Gln Val Phe Ser Glu Glu
165    170    175
Pro Arg Asp Cys Phe Met Leu Glu Thr Thr Tyr Pro Leu Val Gly Phe
180    185    190
Glu Thr Ser Thr Glu Asp Gly Asn Thr Tyr Ala Val Phe Leu Thr Gly
195    200    205
Val Gly Leu Glu Arg Ser Leu Pro Lys Tyr Val Pro Val Phe Asp Met
210    215    220
Asn Ala Gly Ile Gln Thr Leu Asn Met Thr Gly Leu Arg Met Ala Lys
225    230    235    240
Leu Pro Val Leu Cys Met Phe Gly Arg Thr Glu Tyr Asp Asn Leu Glu
245    250    255
Asp Phe Tyr Ile Thr Ser Ile Glu Thr Gln Ser Phe Asp Glu Glu Glu
260    265    270
Asn Asp Ala Arg Met Arg Cys His Thr Glu Asp Leu Glu Arg Lys Lys
275    280    285
Arg Met Asn Asp Ala Pro Ala Ile Thr Pro His Val Ala Val Tyr Asp
290    295    300
Tyr Ser Gly Asp Gly Lys Glu Gln Leu Leu Tyr Met Ile Thr Glu Tyr
305    310    315    320
Glu Asn Thr Ala Ser Trp Cys Asn Ala Asn Gly Val Val Thr Ser Asp
325    330    335

```

Ser	Gly	Phe	Ser	Asn	Glu	Cys	Ala	Ile	Ser	Asp	Met	Asn	Asp	Leu	Cys
			340					345					350		
Cys	Phe	Ala	Asp	Cys	Ile	Asp	Val	Thr	Val	Asn	Asn	Glu	Glu	His	Glu
		355					360					365			
Glu	Arg	Ser	Met	Asn	Ile	Val	Val	Glu	Ser	Asp	Arg	Arg	Leu	Phe	Asp
	370				375					380					
Asp	Ser	Pro	Ile	Lys	Thr	Glu	Glu	Asp	Gly	Glu	Asn	Ser	Ser	Ser	Ser
385				390					395						400
Ser	Ser	Ser	Pro	Thr	Val	Pro	Pro	Pro	Thr	Pro	Tyr	Glu	Gly	Asn	Ala
			405						410					415	
Val	Val	Glu	Gly	Glu	Glu	Glu	Glu	Glu	Ile	Asp	Glu	Asp	Glu	Ser	
		420					425					430			
Ser	Lys	Tyr	Glu	Gly	Ser	Glu	Asp	Ala	Leu	Val	Met	Lys	Lys	Leu	Ala
	435					440						445			
Lys	Leu	Ser	Thr	Met	Lys	Gln	Met	Arg	Arg	Val	Lys	Asn	Glu	Pro	Ala
	450				455					460					
Leu	Lys	Ile	Thr	Ser	Gly	Gly	Asn	Asn	Ser	Ser	Ser	Ser	Ile	Asn	Asn
465				470						475					480
Glu	Asp	Asp	Gly	Asp	Asp	Asp	Asp	Ala	Val	Asp	Ala	Thr	Ala	Leu	Cys
			485						490					495	
Pro	Gln	Thr	Glu	Ala	Thr	Val	Lys	Asn	Ser	Phe	Met	Ala	Pro	Asn	Asp
		500					505						510		
Glu	Arg	Thr	Glu	Asn	Ile	Leu	Tyr	Glu	Thr	Met	Gln	Ile	Ser	Leu	Ala
	515				520						525				
Lys	Ile	Cys	Asn	Asn	Pro	Ser	Ser	Met	Ser	Ser	Tyr	Arg	Val	Phe	Thr
	530				535					540					
Asn	Lys	Leu	Gln	Glu	Cys	Leu	Asn	Thr	Met	Asp	Asp	Ser	Ile	Arg	Arg
545				550						555					560
Arg	Pro	Thr	Ile	Trp	Thr	Glu	Glu	Ser	Gln	Gln	Phe	Ala	Lys	Gly	Leu
			565					570						575	
Leu	Phe	Asp	Glu	Val	Val	Thr	Ser	Ile	Val	Ala	His	Gln	Met	Ala	Gln
		580					585						590		
Asp	Ile	Cys	Lys	Ser	Glu	Ile	Phe	Gly	Gly	Met	Phe	Asn	Ala	Asn	Ser
	595					600						605			
Thr	Asn	Ile	Lys	Gly	Lys	Tyr	Glu	Gly	Gln	Lys	Lys	Ser	Leu	Tyr	Gly
	610				615					620					
Asn	Lys	His	Ser	Ser	Cys	Phe	Lys	Thr	Asn	Thr	Glu	Ser	Asn	Val	Asn
625				630						635					640
Asn	Ala	Leu	Phe	Ala	Trp	Val	Lys	Ser	Lys	Leu	His	Ser	Gly	Thr	Val
			645					650						655	
Ile	Pro	Asn	Val	Phe	Ser	Phe	Lys	Met	Ala	Ser	Glu	Lys	Pro	Ser	Lys
		660						665					670		
Met	Lys	Arg	Lys	Arg	Thr	Ser	Ser	Ala	Ser	Ser	Ser	Asn	Asp	Glu	His
	675				680							685			
Gln	Glu	Pro	Ser	Thr	Lys	Met	Met	Lys	Asn	Asp	Glu	Gly	Glu	Lys	Val
	690				695					700					
Ala	Gln	Glu	Ser	Ser	Ser	Pro	Ser	Ser	Ser	Ser	Thr	Pro	Glu	Gln	Gln
705				710						715					720
Gln	Gln	Ala	Gly	His	Asp	Lys	Glu	Thr	Ile	Asn	Leu	Ile	Pro	Leu	Ser
			725					730						735	
Phe	Ile	Lys	Met	Pro	Arg	Ser	Asn	Val	Asn	Gly	Ser	Ala	Ser	Tyr	Leu
		740					745					750			
Ser	Glu	Ile	Phe	Gly	Gln	Arg	Leu	Cys	Gly	Leu	Ser	Asp	Ala	Ser	Ser
	755						760					765			
Thr	Phe	Lys	Arg	Met	Cys	Lys	Thr	Phe	Glu	Asp	Leu	Glu	Asn	Glu	Ile
	770				775					780					
Met	Arg	Ser	Ser	Phe	Thr	Arg	Leu	Thr	Arg	Tyr	Glu	Arg	Glu	Val	Thr
785				790						795					800
Arg	Leu	Tyr	Glu	Lys	Cys	Arg	Ser	Gln	Ala	Val	Asp	Ile	Glu	Glu	Asn
			805					810					815		
Glu	Met	Asp	Val	Leu	Ser	His	Gln	Gly	Glu	Leu	Phe	Ala	Glu	Phe	Leu

```

      820      825      830
Glu Asp Pro Ile Ala Tyr Phe Glu Glu Val Leu Glu Asn Ile Lys Ser
      835      840      845
Trp Ser Leu Glu Asn Val Asn Thr Pro Lys Arg Lys Asn Lys Tyr Ala
      850      855      860
Lys Val Leu Val Ser Val Asn Ala Ile Arg Arg Thr Tyr Glu Glu Tyr
      865      870      875      880
His Ala Phe Ser Lys Phe Val Pro Met Phe Leu Phe Asn Leu Ile Lys
      885      890      895
Arg Glu Leu Glu Gly Asp Asn Tyr Thr His Asp Val His Phe Ser Ser
      900      905      910
Thr Cys Leu Trp Tyr Leu Thr Val Met Thr Arg Asn Arg Ile Cys Asp
      915      920      925
Val Leu Gln Tyr Ile Asn Asn Asn Asn Asn Asp Asn Glu Glu Thr Asp
      930      935      940
Ile Val Glu Glu Glu Glu Glu Gly Glu Gly Glu Asp Lys Met Glu
      945      950      955      960
Glu Ser Met Asp Val Glu Gln Gln Lys Gln Val Arg Lys Gly Gly Arg
      965      970      975
Lys Lys Gly Gln Lys Phe Asn Ser Ile Gly Asp Gln Val Ile Arg Lys
      980      985      990
Phe Val Lys Ser Leu Cys Glu Asn Ser Met Val Val Ser Ile Ala Ile
      995      1000      1005
Asn Ser Leu Ile Ser Gly Ile Ser Trp Met Asn Lys Lys Ile Pro Pro
      1010      1015      1020
Gly Phe Leu Lys Asp Ser Ser Thr Ile Asn Thr Leu Asp Glu Val Ser
      1025      1030      1035      1040
Arg Phe Val Phe Ser Asp Val Lys Ile Asn Arg Lys Ile Asn Gly Thr
      1045      1050      1055
Asp Asp Lys Tyr Glu Thr Val Phe Gly Val Ser Thr Arg Val Asp Ser
      1060      1065      1070
His Ile Val Gly Pro Phe Ser Ile Pro Val Asp Phe Ser Ser Ala Gly
      1075      1080      1085
Leu Asp Lys Ala Ser Cys Gly Lys Leu Tyr Val Asn Thr Ile Asp Gly
      1090      1095      1100
Lys Gly Ile Leu Thr Ile Ser Pro Lys Tyr Asp Ser Leu Asn Asp Glu
      1105      1110      1115      1120
Asp Val Asp Ser Thr Thr Thr Asp Lys Leu Glu Lys Asp Ile Leu His
      1125      1130      1135
Leu Ser Lys His Asp Thr Phe Phe Asn Ile Asn Lys Asn Lys Val Leu
      1140      1145      1150
Pro Phe Tyr Asn Ile Ser Pro Ser Ser Ser Leu Thr Glu Lys Lys Lys
      1155      1160      1165
Thr Lys Phe Asn Arg Lys Lys Ile Ser Ser Gly Met Ser Asn Asn Asn
      1170      1175      1180
Gly Met Cys Val Gln Thr Pro Ser Ser Ser Asn Ser Val Ser Ser Val
      1185      1190      1195      1200
Ser Ser Ile Val Ala Pro Ser Ser Ser Val Leu Ala Leu Ser Cys Ser
      1205      1210      1215
Leu Ser Ser Thr Lys Lys Lys Ser Ile Trp Asn Glu Asn Met Phe Leu
      1220      1225      1230
Thr Ser Arg Asn Met Trp Arg Cys Gly Phe Val Val Pro Pro Lys Leu
      1235      1240      1245
Cys Ser Phe Ile Val Asn His Arg His Ala Val Lys Leu Val Ala Glu
      1250      1255      1260
Thr Ala Pro Lys Thr Lys Leu Cys Arg Asn Ile Ile Asp Arg Asn Arg
      1265      1270      1275      1280
Lys Ile Arg Phe Asn Gly Leu Lys Lys Val Cys Lys Ser Val Ser Ala
      1285      1290      1295
Phe Thr Gly Glu Ser Thr Tyr Leu Leu Asn Lys Asn Met Thr Ala Thr
      1300      1305      1310

```

Ser Pro Ser Asp Leu Asn Leu Cys Ile Tyr Thr Ser Ser Leu Asn Asp  
 1315 1320 1325  
 Pro Leu Tyr Thr Cys Lys Leu Thr His Glu Glu Tyr Gln Asp Gly Asn  
 1330 1335 1340  
 Ala Leu Asp Asp Tyr Gly Ala Val Phe Val Asn Tyr Thr Phe Lys Ser  
 1345 1350 1355 1360  
 Ile Lys Ser Cys Ser Ser Lys Asp Glu Thr Ala Asp Asp Asn Ala Ala  
 1365 1370 1375  
 Ala Ala Asp Asp Asp Gly Ser Thr Thr Ser Thr Ser Ser Thr Asp  
 1380 1385 1390  
 Thr Asp Ala Ala Ala Ile Gln Asp Phe Met His Val Met Ile Lys Lys  
 1395 1400 1405  
 Ile Asp Ala Met Lys Asp Ile Arg Gly Lys Tyr Lys Lys Ser Leu Ala  
 1410 1415 1420  
 Lys Lys Thr Lys Lys His  
 1425 1430

&lt;210&gt; 242

&lt;211&gt; 909

&lt;212&gt; DNA

&lt;213&gt; SHRIMP

&lt;400&gt; 242

atggtgtcgt ctattaccca cctctctctg ttgttcgtcg tggctgtagt agcttccgctc 60  
 gtttttacaa ctgaaggagc tagtgtgaga gtgaaacggt gtgctgttag cccgtgcccc 120  
 gacgttattg accccgacca ccgctgccaa ggcgcactgt gccgcaggtc tactcgagga 180  
 ggtgacgacg acgacgacga tgacgatgga ggaactttcg atacagtagg gtctgggtata 240  
 cttggacgca aaaagcgtgc cgcacctcca cctgaggatg aagaagagga tgatttctac 300  
 cgcaaaaagc gtgccgcacc tccacctgag gatgaagaag aggatgattt ctaccgcaaa 360  
 aagcgtgccg cacctccacc tgaggatgaa gaagaggatg agttctaccg caaaaagcgt 420  
 gccgcacctc cacctgagga tgaagaagag gatgagttct accgcaaaaa gcgtgccgca 480  
 cctccacctg aggatgaaga agaggatgag ttctaccgca aaaagcgtgc cgcacctcca 540  
 cctgaggatg aagaagagga tgagttctac cgcaaaaagc gtgccgcacc tccacctgag 600  
 gatgaagaag aggatgagtt ctaccgcaaa aagcgtgccg cacctccacc tgaggatgaa 660  
 gaagaggatg agttctaccg caaaaagcgt gccgcacctc cacctgagga tgaagaagag 720  
 gatgatttct accgcaaaaa gcgtgccgca cctccacctg aggatgaaga agaggatgat 780  
 ttctaccgca aaaagcgtgc cgcacctcca cctgaggatg aagaagagga tgatttctac 840  
 cgcaaaaagc gtgccgcacc tccacctgag gatgaagaag aggatgattt ctaccgcaaa 900  
 aagcgtaa 909

&lt;210&gt; 243

&lt;211&gt; 302

&lt;212&gt; PRT

&lt;213&gt; SHRIMP

&lt;400&gt; 243

Met Val Ser Ser Ile Thr His Leu Ser Leu Leu Phe Val Val Ala Val  
 1 5 10 15  
 Val Ala Ser Val Val Phe Thr Thr Glu Gly Ala Ser Val Arg Val Lys  
 20 25 30  
 Arg Cys Ala Val Ser Pro Cys Pro Asp Val Ile Asp Pro Asp His Arg  
 35 40 45  
 Cys Gln Gly Arg Leu Cys Arg Arg Ser Thr Arg Gly Gly Asp Asp Asp  
 50 55 60  
 Asp Asp Asp Asp Asp Gly Gly Thr Phe Asp Thr Val Gly Ser Gly Ile  
 65 70 75 80  
 Leu Gly Arg Lys Lys Arg Ala Ala Pro Pro Glu Asp Glu Glu Glu  
 85 90 95  
 Asp Asp Phe Tyr Arg Lys Lys Arg Ala Ala Pro Pro Pro Glu Asp Glu  
 100 105 110

WO 01/38351

PCT/US00/28888

340

Glu Glu Asp Asp Phe Tyr Arg Lys Lys Arg Ala Ala Pro Pro Pro Glu  
 115 120 125  
 Asp Glu Glu Glu Asp Glu Phe Tyr Arg Lys Lys Arg Ala Ala Pro Pro  
 130 135 140  
 Pro Glu Asp Glu Glu Glu Asp Glu Phe Tyr Arg Lys Lys Arg Ala Ala  
 145 150 155 160  
 Pro Pro Pro Glu Asp Glu Glu Glu Asp Glu Phe Tyr Arg Lys Lys Arg  
 165 170 175  
 Ala Ala Pro Pro Pro Glu Asp Glu Glu Asp Glu Phe Tyr Arg Lys  
 180 185 190  
 Lys Arg Ala Ala Pro Pro Pro Glu Asp Glu Glu Glu Asp Glu Phe Tyr  
 195 200 205  
 Arg Lys Lys Arg Ala Ala Pro Pro Pro Glu Asp Glu Glu Glu Asp Glu  
 210 215 220  
 Phe Tyr Arg Lys Lys Arg Ala Ala Pro Pro Pro Glu Asp Glu Glu Glu  
 225 230 235 240  
 Asp Asp Phe Tyr Arg Lys Lys Arg Ala Ala Pro Pro Pro Glu Asp Glu  
 245 250 255  
 Glu Glu Asp Asp Phe Tyr Arg Lys Lys Arg Ala Ala Pro Pro Pro Glu  
 260 265 270  
 Asp Glu Glu Asp Asp Phe Tyr Arg Lys Lys Arg Ala Ala Pro Pro  
 275 280 285  
 Pro Glu Asp Glu Glu Glu Asp Asp Phe Tyr Arg Lys Lys Arg  
 290 295 300

<210> 244  
 <211> 1119  
 <212> DNA  
 <213> SHRIMP

<400> 244  
 atggataact tgaagggga atttgttgcg cttaaaacag acctcaccca ttacaaaaca 60  
 cagtttgata gatctatatt ggtatattgtt gatgtgtgtg gtagattata tggtatagta 120  
 aatagtgaac aaacagctaa aaaggaaggt ctacgaacta gagtggaaca gcaagccaca 180  
 gagatacaac aattcaagga cgaaataaac aacaaatata atgctctaac aaatactttg 240  
 gatgatata tctacatttt tgatcatgga gggagtgttca aaagagcaaa acataaggcc 300  
 ataattgaag cgagggaata ctctaaaccg ctgagggaat tagagtgcac gtttacgcgt 360  
 atagcggaca tgtaaacctt gacttttatg actgtgtaca ccaatatcat tactgaattt 420  
 agacactcta gtgaacaagc cactaatagt ataatgtca ccctcggacg tcttttcttg 480  
 tgtgacgact tgtgcaatca attacaaaaa gaagaggaag aagaggaaga tttgaaacag 540  
 aaattcatta ctttccatgc gaacctatac atgctggaca cagcctaata gaaagatttg 600  
 ataattttca aagatgtcat acaacaactt cactgtgatt tgcaaaaagga tacctatgct 660  
 gtaaaagaag gtgtggccat tagatgtgag aaacagatga acgaaataag tcaatacagg 720  
 gacaacctca aggataatta caatacattt tcaaacattt tgaatgaaat tgtctacatt 780  
 tttgatcacg ggggacattt tgaagaagta aaacacaaaag ccataactct gactagaaat 840  
 tacttgaaaa cactcatggg attaaaatgc atgttcaaac gcataatccga aatgttgtca 900  
 ttgacttttc taacagtgt cactaatgtt atagcagaat ttataaacgc tagcaatatt 960  
 tctgatagag agatcaataa ttatcttgtc caacttgtaa catgtaacga attgtgcaac 1020  
 caactcccca aacctaaaca ataccgtccc ctcagtttga tagataacat agcttatttt 1080  
 tctctttctg tccaaaaaca tctgagtggg tttcttttag 1119

<210> 245  
 <211> 368  
 <212> PRT  
 <213> SHRIMP

<400> 245  
 Met Asp Asn Leu Lys Gly Glu Phe Val Lys Thr Asp Leu Thr His Tyr  
 1 5 10 15  
 Lys Thr Gln Leu Asp Arg Ser Ile Leu Val Phe Val Asp Val Val Gly



			20					25					30		
Arg	Leu	Tyr	Val	Ile	Val	Asn	Ser	Glu	Gln	Thr	Ala	Lys	Lys	Glu	Gly
		35					40					45			
Leu	Ala	Thr	Arg	Val	Ala	Lys	Gln	Ala	Thr	Glu	Ile	Gln	Gln	Phe	Lys
		50				55					60				
Asp	Glu	Ile	Asn	Asn	Lys	Tyr	Asn	Ala	Leu	Thr	Asn	Thr	Leu	Asp	Asp
65					70					75				80	
Ile	Ile	Tyr	Ile	Phe	Asp	His	Gly	Gly	Ser	Phe	Lys	Arg	Ala	Lys	His
				85					90					95	
Lys	Ala	Ile	Ile	Glu	Ala	Arg	Glu	Tyr	Ser	Lys	Pro	Leu	Arg	Glu	Leu
			100					105					110		
Glu	Cys	Met	Phe	Thr	Arg	Ile	Ala	Asp	Met	Leu	Thr	Leu	Thr	Phe	Met
		115					120					125			
Thr	Val	Tyr	Thr	Asn	Ile	Ile	Thr	Glu	Phe	Arg	His	Ser	Ser	Glu	Gln
		130			135						140				
Ala	Thr	Asn	Ser	Ile	Asn	Val	Thr	Leu	Gly	Arg	Leu	Phe	Leu	Cys	Asp
145				150						155				160	
Asp	Leu	Cys	Asn	Gln	Leu	Pro	Lys	Glu	Glu	Glu	Glu	Glu	Glu	Asp	Leu
			165					170						175	
Lys	Gln	Lys	Phe	Ile	Thr	Phe	His	Ala	Asn	Leu	Tyr	Met	Leu	Asp	Thr
			180					185					190		
Arg	Leu	Lys	Lys	Asp	Leu	Ile	Ile	Phe	Lys	Asp	Val	Ile	Gln	Gln	Leu
		195				200						205			
His	Val	Ile	Leu	Gln	Lys	Asp	Thr	Tyr	Ala	Val	Lys	Glu	Gly	Val	Ala
		210				215					220				
Ile	Arg	Cys	Ala	Lys	Gln	Met	Asn	Glu	Ile	Ser	Gln	Tyr	Arg	Asp	Asn
225				230						235				240	
Leu	Lys	Asp	Asn	Tyr	Asn	Thr	Phe	Ser	Asn	Ile	Leu	Asn	Glu	Ile	Val
			245					250					255		
Tyr	Ile	Phe	Asp	His	Gly	Gly	His	Phe	Glu	Glu	Val	Lys	His	Lys	Ala
			260					265					270		
Ile	Thr	Leu	Thr	Arg	Asn	Tyr	Leu	Lys	Thr	Leu	Met	Gly	Leu	Lys	Cys
		275					280					285			
Met	Phe	Lys	Arg	Ile	Ser	Glu	Met	Leu	Ser	Leu	Thr	Phe	Leu	Thr	Val
		290				295					300				
Tyr	Thr	Asn	Val	Ile	Ala	Glu	Phe	Ile	Asn	Asn	Ile	Ser	Asp	Arg	Glu
305				310						315				320	
Ile	Asn	Asn	Tyr	Leu	Val	Gln	Leu	Val	Thr	Cys	Asn	Glu	Leu	Cys	Asn
			325					330					335		
Gln	Leu	Pro	Lys	Pro	Lys	Gln	Tyr	Arg	Pro	Leu	Ser	Leu	Ile	Asp	Asn
			340					345					350		
Ile	Ala	Tyr	Phe	Ser	Leu	Ser	Val	Gln	Lys	His	Leu	Ser	Gly	Phe	Leu
		355					360					365			

<210>	246
<211>	1545
<212>	DNA
<213>	SHRIMP

<400> 246						
atggatagtt	gttgccctgat	atcgaggata	acaccagaac	tggctggcaa	gttaacctgg	60
atcttcatac	cagaaaacaa	cttcaagatt	gtccagaact	cactcccaga	cgaccaagtt	120
atctcccaat	tcagatat	cgaccataga	cattgctata	cgtttatgga	gattttgatg	180
gcaacacatta	aaatccaaga	caggaatacaa	aacaccacag	ccatatgtga	attgacaact	240
gcgaagagaag	gacttttatg	taggagaacc	atacctgtat	ttttgggttc	agaggaaaaa	300
cgagaagagt	tattggggaa	tctccctgaa	ggtgcagaaa	ttttcaggcc	tagagaagtt	360
atgcaagtaa	ttggtactct	cttggacaag	aaactagaaa	ttgacgacgg	tatagcttct	420
gtaaaggctg	ccctctgtgc	tggttcatca	tcgttatacc	taatcatgag	ccacatagtg	480
aaaatgacct	tttctgctat	cacaacaatg	aaggatataa	acgaagaata	tttcgtagag	540
tttatatttc	qtcataaaca	attcctcaac	cctqaattct	tcaaqcacct	tatatctttg	600

```

ctcaagaatt ccaggaagga acatgttgcc catctagtaa gacgtctaga acactttctc 660
atgctatgga cccitttccaa gatgaggttc acagaaatgg aagaaaacta cttcccaatc 720
tccagcgata gtgattacgg catctgtgaa aaatgtgcac gaaaaactcc caaatacaag 780
ctccgtatatt ttagggaacg aaaatgctgc gatagatgtt gccgtcttta tcaccaacaa 840
ccgcctccgg aggtgtataa ttgggatgga aaaataaccc aacaatccaa taaaggctac 900
attaatgcag gcgatgaaat tatcggcacg ctaaaactcaa atgataaggg aaaaacattc 960
cctcctatac ctaagatggt tgtacgaaga gtggtggacg gtgtctacgg gcaaggaact 1020
atcctgtcaa agattttgaa gttcaggcag gcaaataatcc ccacgtgtct attcgtgaca 1080
tgcaataaat gcaataggat tttcaggctc actatcttag ggcctacaag aaacatcctt 1140
tgcccacctt gcagaaagaa aagtgttgca gtaaatacac aacagaaagg agaaaataaa 1200
ccttcgtttg tgcaaaaagg aacaaaacgt ctacgagtgg ataccggtag caacaagaac 1260
acgttagaaa aattctgttc ctgggaaaga ttcaatactg aagttttgct cccttggtct 1320
ggctacacta ttgagtctaa gtggcagaac tgggaatctt ttctgggtta ttcgagtacc 1380
agatataagg aactgtgggc ctttgtgaac aaacaggaaa tatcttccat gaaagactcc 1440
tacataaaaa ttgaagacat cgaccagtta ttgaggagta tcttgcaaga ccagaagggt 1500
gtatttgaga ccgtctgcaa aataaagagc agagatggtt tgtga 1545

```

&lt;210&gt; 247

&lt;211&gt; 514

&lt;212&gt; PRT

&lt;213&gt; SHRIMP

&lt;400&gt; 247

```

Met Asp Ser Cys Cys Leu Ile Ser Arg Ile Thr Pro Glu Leu Ala Gly
 1          5          10          15
Lys Leu Thr Trp Ile Phe Ile Pro Glu Asn Asn Phe Lys Ile Val Gln
 20          25          30
Asn Ser Leu Pro Asp Asp Gln Val Ile Ser Gln Phe Arg Tyr Phe Asp
 35          40          45
His Arg His Cys Tyr Thr Phe Met Glu Ile Leu Met Ala Asn Ile Lys
 50          55          60
Ile Gln Asp Arg Lys Gln Asn Thr Thr Ala Ile Cys Glu Leu Thr Thr
 65          70          75          80
Gly Arg Glu Gly Leu Leu Cys Arg Arg Thr Ile Pro Val Phe Leu Gly
 85          90          95
Ser Glu Glu Lys Arg Glu Glu Leu Leu Gly Asn Leu Pro Glu Gly Ala
100          105          110
Glu Ile Phe Arg Pro Arg Glu Val Met Gln Val Ile Gly Thr Leu Leu
115          120          125
Asp Lys Lys Leu Glu Ile Asp Asp Gly Ile Ala Ser Val Lys Ala Ala
130          135          140
Leu Cys Ala Gly Ser Ser Ser Leu Tyr Leu Ile Met Ser His Ile Val
145          150          155          160
Lys Met Thr Phe Ser Ala Ile Thr Asn Met Lys Asp Ile Asn Glu Glu
165          170          175
Tyr Phe Val Asp Phe Ile Phe Arg His Lys Gln Phe Leu Asn Pro Glu
180          185          190
Phe Phe Lys His Leu Ile Ser Leu Leu Lys Asn Ser Arg Lys Glu His
195          200          205
Val Ala His Leu Val Arg Arg Leu Glu His Phe Leu Met Leu Trp Thr
210          215          220
Leu Ser Lys Met Arg Phe Thr Glu Met Glu Glu Asn Tyr Phe Pro Ile
225          230          235          240
Ser Ser Asp Ser Asp Tyr Gly Ile Cys Glu Lys Cys Ala Arg Lys Thr
245          250          255
Pro Lys Tyr Lys Leu Arg Ile Phe Arg Glu Arg Lys Cys Cys Asp Arg
260          265          270
Cys Cys Arg Leu Tyr His Gln Gln Pro Pro Pro Glu Val Tyr Asn Trp
275          280          285
Asp Gly Lys Ile Thr Gln Gln Ser Asn Lys Gly Tyr Ile Asn Ala Gly
290          295          300

```

[illegible]

<210>	248
<211>	1242
<212>	DNA
<213>	SHRIMP

<400>	248						
atggagtgcaa	tcaaaactgtt	caccgttgct	ggtctgaata	tggagcaagc	caaccaagtg	60	
gctgaagaaa	tcaagtcaga	atataaaacc	gaggaggaaa	agaggattgc	ccaggaagtg	120	
tttgacaaat	tcaccaaaaa	actcattatg	caagtagata	cgtctaaaca	cttacttaca	180	
agagaaaaac	ccaaccgttt	tgtatccgc	ccatttgttc	atgaagatct	ctgggaaatg	240	
tacaaaaaag	agggtgcctg	tttttggaca	ttggaagaga	ttgatttcga	aagggatcct	300	
aaagattggg	agaaactcac	tcaagatgag	aaggatttca	ttctccagat	tctggcgttc	360	
tttgcatcct	ctgacggaat	tgtaattgaa	aatcttaca	cacgtcttcg	tcaagtggcg	420	
cagattccag	aagcgaggag	tttctttgac	ttccaagttg	gaatggagag	tattctatggc	480	
aacgtctacg	gagaactgat	tgatagactg	gtgcccgcag	aaaaagacaa	ggctatcttg	540	
tttaacgctg	cacaacactt	ccccgcacatc	aagaagaagg	agcagtgggc	tattaattgg	600	
atgcaaagca	ataacgattt	ggcggaacta	attgttgctt	ttgctgcagt	tgaaggaatc	660	
ttcttttagtg	gtgcattcgc	atccattttc	tggatcaaga	acaggggtat	tttgcttggt	720	
ctcacctcct	ccaatgagtt	catttctag	gacgaaggtc	ttcatcgca	ctttgcatgc	780	
atgctgttg	aaaagggttt	tgttgatacc	ccatcaagag	aaaggattct	tgaattgtc	840	
actgaagcgc	tccgaattga	acaagaattt	ctcacagttt	ccctgcctgt	taaattagtg	900	
ggaatgaact	gcaagttgat	gagccagtac	attgaatttg	tggcagata	actattgggt	960	
gaaatgggac	tagaaaagca	ctataatgtt	accaaccctt	tcccatcatt	ggacaatatt	1020	
tccctcgaga	ataagaccaa	cttttttgaa	aagagagtcg	ccgagtatca	acgtgccag	1080	
gtcatggctt	ctatcaataa	gatcaagaag	gaccaacaaa	ccaagaaac	tggttctcct	1140	
ctcccaattc	tgactgcacc	tctccagtc	tcttctctat	catccgaaca	agaagatgtt	1200	
gaagacggcg	tcggggacta	catcagttat	gacqattttt	aq		1242	

```
<210> 249
<211> 409
<212> PRT
```

&lt;213&gt; SHRIMP

&lt;400&gt; 249

```

Met Glu Ser Ile Lys Leu Phe Thr Val Ala Gln Met Glu Gln Ala Asn
 1      5      10      15
Gln Val Ala Glu Glu Ile Lys Ser Glu Tyr Lys Thr Glu Glu Glu Lys
 20      25      30
Arg Ile Ala Gln Glu Val Phe Asp Lys Phe Thr Lys Lys Leu Ile Met
 35      40      45
Gln Val Asp Thr Ser Lys His Leu Leu Thr Arg Glu Asn Pro Asn Arg
 50      55      60
Phe Val Ser Arg Pro Ile Val His Glu Asp Leu Trp Glu Met Tyr Lys
 65      70      75      80
Lys Glu Val Ala Cys Phe Trp Thr Leu Glu Glu Ile Asp Phe Glu Arg
 85      90      95
Asp Pro Lys Asp Trp Glu Lys Leu Thr Gln Asp Glu Lys Asp Phe Ile
100     105     110
Leu Gln Ile Leu Ala Phe Phe Ala Ser Ser Asp Gly Ile Val Ile Glu
115     120     125
Asn Leu Thr Thr Arg Leu Arg Gln Val Ala Gln Ile Pro Glu Ala Arg
130     135     140
Ser Phe Phe Asp Phe Gln Val Gly Met Glu Ser Ile His Gly Asn Val
145     150     155     160
Tyr Gly Glu Leu Ile Asp Arg Leu Val Pro Asp Glu Lys Asp Lys Ala
165     170     175
Ile Leu Phe Asn Ala Ala Gln His Phe Pro Ala Ile Lys Lys Lys Glu
180     185     190
Gln Trp Ala Ile Asn Trp Met Gln Ser Asn Asn Asp Leu Ala Glu Leu
195     200     205
Ile Val Ala Phe Ala Ala Val Glu Gly Ile Phe Phe Ser Gly Ala Phe
210     215     220
Ala Ser Ile Phe Trp Ile Lys Asn Arg Gly Ile Leu Pro Gly Leu Thr
225     230     235     240
Ser Ser Asn Glu Phe Ile Ser Arg Asp Glu Gly Leu His Arg Asp Phe
245     250     255
Ala Cys Met Leu Leu Lys Lys Gly Phe Val Asp Thr Pro Ser Arg Glu
260     265     270
Arg Ile Ile Val Thr Glu Ala Val Arg Ile Glu Gln Glu Phe Leu Thr
275     280     285
Val Ser Leu Pro Val Lys Leu Val Gly Met Asn Cys Lys Leu Met Ser
290     295     300
Gln Tyr Ile Glu Phe Val Ala Asp Lys Leu Leu Val Glu Met Gly Leu
305     310     315     320
Glu Lys His Tyr Asn Val Thr Asn Pro Phe Pro Phe Met Asp Asn Ile
325     330     335
Ser Leu Glu Asn Lys Thr Asn Phe Phe Glu Lys Arg Val Ala Glu Tyr
340     345     350
Gln Arg Ala Gln Val Met Ala Ser Ile Asn Lys Ile Lys Lys Asp Gln
355     360     365
Gln Thr Gln Glu Thr Gly Ser Pro Leu Pro Ile Leu Thr Ala Pro Pro
370     375     380
Pro Val Ser Ser Ser Ser Ser Glu Gln Glu Asp Val Glu Asp Gly Val
385     390     395     400
Gly Asp Tyr Ile Ser Tyr Asp Asp Phe
405

```

&lt;210&gt; 250

&lt;211&gt; 915

&lt;212&gt; DNA

&lt;213&gt; SHRIMP

&lt;400&gt; 250

```

atgaatttac ttccaatatt cctgacaacc ttttttgttg cggtagatgc atgctcttgc 60
tctaccatct gcctttttacc tgatgggaag aaacaaccct tggtttttga ttctgtatta 120
gaagaggtgg tataccctac agatgtgtgt gggccaaagg gagctggcga attattcact 180
ggtgtggatc ttttgaccct ctgtatagga ggtaaaaaca atggaggtga atggtcagga 240
aaaggtcctt gtccaaggat caataacgct gtcgttgaac gagattactc ccttgacgag 300
gaggattgta aagggtttag aaaggggttc cgaattcctg gcactgacca ttttcatact 360
gtcttttccc tttgttgggt agacagagat atgcacgcca agtgggtgcg caacaaaata 420
aaccctggta tagtaactga tgatgaagat ttggtagatt ctggtattag gactaaattt 480
aaatactctt ctaaaatttt tggtaaagga ttcaatccga gacctcttta ctccctcgac 540
tatcaagaga ggattaagat attaaagtct cattttaaca agaggacggg taatttcttt 600
gctcgaggcc acttggctcc ggctggagat tttttcctcg cttcagagag atgggcaact 660
tttgctctag agaatgcagt acctcaaata cagaaccata acaatggtga atggaaagat 720
attgaaaata gtgcaagaac tacgccaggt gccgcgtggg ctgagactgg accaatattt 780
taccaacaca agaagaagga atatctagac aagaagaaga agtacatccc tatccctcat 840
gccctctaca agattgtgta cgacaagaat aacaaggaat tgttccgtgt acagagtgat 900
atgtcttgga aataa 915

```

&lt;210&gt; 251

&lt;211&gt; 298

&lt;212&gt; PRT

&lt;213&gt; SHRIMP

&lt;400&gt; 251

```

Met Asn Leu Leu Pro Ile Phe Leu Thr Thr Phe Phe Val Ala Val Asp
1      5      10
Ala Cys Ser Cys Ser Thr Ile Cys Leu Leu Pro Asp Gly Lys Lys Gln
20     25     30
Pro Leu Val Phe Asp Ser Val Leu Glu Glu Val Val Tyr Pro Thr Asp
35     40     45
Val Cys Gly Pro Lys Gly Ala Gly Glu Leu Phe Thr Gly Val Asp Leu
50     55     60
Leu Thr Leu Cys Ile Gly Gly Lys Asn Asn Gly Gly Glu Trp Ser Gly
65     70     75
Lys Gly Pro Cys Pro Arg Ile Asn Asn Ala Val Val Glu Arg Asp Tyr
85     90     95
Ser Leu Asp Glu Glu Asp Cys Lys Gly Phe Arg Lys Gly Phe Arg Ile
100    105    110
Pro Gly Thr Asp His Phe His Thr Val Phe Ser Leu Cys Trp Val Asp
115    120    125
Arg Asp Met His Ala Lys Trp Val Arg Asn Lys Ile Asn Pro Gly Ile
130    135    140
Val Thr Asp Asp Glu Asp Leu Val Asp Ser Gly Ile Arg Thr Lys Phe
145    150    155
Lys Tyr Ser Ser Lys Ile Phe Gly Lys Gly Phe Asn Pro Arg Pro Lys
165    170    175
Leu Asp Tyr Gln Glu Arg Ile Lys Ile Leu Lys Ser His Phe Asn Lys
180    185    190
Arg Thr Gly Asn Phe Phe Arg His Leu Ala Pro Ala Gly Asp Phe Phe
195    200    205
Leu Ala Ser Trp Ala Thr Phe Ala Leu Glu Asn Ala Val Pro Gln Ile
210    215    220
Gln Asn His Asn Asn Gly Glu Trp Lys Asp Ile Glu Asn Arg Ala Arg
225    230    235
Thr Thr Pro Gly Ala Ala Trp Ala Glu Thr Gly Pro Ile Phe Tyr Gln
245    250    255
His Lys Lys Lys Glu Tyr Leu Asp Lys Lys Lys Tyr Ile Pro Ile
260    265    270
Pro His Ala Leu Tyr Lys Ile Val Tyr Asp Lys Asn Asn Lys Glu Leu
275    280    285

```

Phe Arg Val Gln Ser Asp Met Ser Trp Lys  
290 295

<210> 252  
<211> 789  
<212> DNA  
<213> SHRIMP

<400> 252  
atgtcgttgg ctgtgacaga agattacggg cacaatgaaa agttgatcaa acggttacaa 60  
acctctgtat atcacacccc tctattaggt gcagaccatg taatgaaatc catatcagac 120  
tacataattt ctctgtcgatt catgaactac acaaatttat taaaacaagt tgaatatgtt 180  
ttcgtatgaag aaacaggagc agttatagct aatatctgtc tgttaaaaaat cctagaaaaga 240  
tgcgcacaga aaggaggaat atatgatgca ccagaagatg ttgcattctt caattctaag 300  
atgggggaag taacgcgcct atttactatt ataggaggta ggcccaatat gacggtgcgg 360  
gttaattttta aacatgggca gacaaataat cctgcctatg gttatctcac agatgataat 420  
gatactacta ctgttactcc tctgtttact cctcctccat ctccagctgc aagaagatcc 480  
ccttttttca cacgactct catatccgag tcgtcttcag ttgaccatta tgtattgatg 540  
catgataacc caaaaagatc ttcattttaag gtgtatgata ttcacgcaga aacctttccc 600  
cataaagctc cttctgttcc taccttcccc cctaaaacct cgtttgaaat ttctgacgtg 660  
actctcgatt gttcaatgga gattttttca cgagacaggg atgtttttaga caatgttcac 720  
gactatatgg ctaacgaccc cgtaccattt ttagtggtatg ttgtgcaccg tggatctagt 780  
ctccgttga 789

<210> 253  
<211> 260  
<212> PRT  
<213> SHRIMP

<400> 253  
Met Ser Leu Ala Val Thr Glu Asp Tyr Gly His Asn Glu Lys Leu Ile  
1 5 10 15  
Lys Arg Leu Gln Thr Ser Val Tyr His Thr Pro Leu Leu Gly Ala Asp  
20 25 30  
His Val Met Lys Ser Ile Ser Asp Tyr Ile Ile Ser Arg Arg Phe Met  
35 40 45  
Asn Tyr Thr Asn Leu Leu Lys Gln Val Glu Tyr Val Phe Asp Glu Glu  
50 55 60  
Thr Gly Ala Val Ile Ala Asn Ile Cys Leu Leu Lys Ile Arg Cys Ala  
65 70 75 80  
Gln Lys Gly Gly Ile Tyr Asp Ala Pro Glu Asp Val Ala Phe Phe Asn  
85 90 95  
Ser Lys Met Gly Glu Val Thr Arg Leu Phe Thr Ile Ile Gly Gly Arg  
100 105 110  
Pro Asn Met Thr Val Arg Val Asn Phe Lys His Gly Gln Thr Asn Asn  
115 120 125  
Pro Ala Tyr Gly Tyr Leu Thr Asp Asp Asn Asp Thr Thr Thr Val Thr  
130 135 140  
Pro Pro Val Thr Pro Pro Pro Ser Pro Ala Ala Arg Arg Ser Pro Phe  
145 150 155 160  
Phe Thr Arg Thr Leu Ile Ser Glu Ser Ser Val Asp His Tyr Val  
165 170 175  
Leu Met His Asp Asn Pro Lys Arg Ser Ser Phe Lys Val Tyr Asp Ile  
180 185 190  
His Ala Glu Thr Phe Pro His Lys Ala Pro Ser Val Pro Thr Phe Pro  
195 200 205  
Pro Lys Thr Ser Phe Glu Ile Ser Asp Val Thr Leu Asp Cys Ser Met  
210 215 220  
Glu Ile Phe Ser Arg Asp Arg Asp Val Leu Asp Asn Val His Asp Tyr  
225 230 235 240

WO 01/38351

347

PCT/US00/28888

Ile Ala Asn Asp Pro Val Pro Phe Leu Val Asp Val Val His Arg Gly  
 245 250 255  
 Ser Ser Leu Arg  
 260

<210> 254  
 <211> 2571  
 <212> DNA  
 <213> SHRIMP

<400> 254  
 atggcatcca cttcttcac c aactaagaaa agagtacatg aagaagatga aaatctcatc 60  
 ccacaaccca aaaagaaaaa atcaaagaaa gtactaccat ttcctgttga caagtataga 120  
 gctgtggata aaaaggtggg aaatctcata cacaagatat tagatcaaga aaaggaccac 180  
 ctttctagta ccgaactgca aatgataact gaatgtaatg gtgcgcgaga agatctgctt 240  
 aaacatcttc tagacgaagg agaatttaac cctactataa ttgaagtagt atcatccatg 300  
 cctattgaaa caatatacga aatactctct tcttctgctg acgacaagaa gttgtgacag 360  
 atatcattat caatgttgat ccacatactt ttcttctgctg ataagggtac tatgtgggta 420  
 tccaacgcgt gcgttcaaaa tgttttgggg aacgactata aagtgggaatt tgaaaatata 480  
 cgtaaaaagt atctgatatt ggaagactta ctgaacggcg tttcaaatca ttggtctgaa 540  
 catggtcctc tttctcacat gctccattct tcaatcccta ttgtacaaga catgttattg 600  
 aacaggctgg tgcgttactt tagcacgtat gatggagatg ctcaattcga tatatcattc 660  
 ataattaata gtgtcttgtg ggaattgat aaaagtgttc tcaacgaatt gacacaattg 720  
 atatcgaggg gtgttttcat tgtgtcgtac gtaccgatgc gtgtacgaac accttcaaa 780  
 gacagtaatc ggccacaaaa tactccttca caaaatatgt cagcactagg tatgaaactc 840  
 aatacatttt catccagaat ctcaagtgtac agaaacaata cctttaaaaa actaaccgag 900  
 ttagtgcata actttgatta cggttccaaa gatgcacat catcatctcc tctcctcct 960  
 tcattatcgg acagcgtcaa cacttttgtg aggttgtaca ccaactatga catattctta 1020  
 aaggtgattt ccgactggaa aatgccttat ggttcttta agaaaacttt tgacgtcctt 1080  
 tattctaagg ggttgatgac attatcagtg tctgaatata cactcaaaaa agagttgggt 1140  
 acgtttttgc gcgccttgaa ggaaagggaa attttaatct ataaaatgga gaagagagac 1200  
 attatatgta tactgaaaaa gtctttgttc ggatttaatt tcaggtgttt aaaacaatta 1260  
 ctccctctct tcaaacactt tttaaaaatt gaagagggtta aacatatagc acgttttgtc 1320  
 tttagagatt acagtctcat gtgcaaaact caaaaagatt tgcagagttt ccctgccata 1380  
 cagtctgctt cacttttcat ggaagaattc ccttggcttg caaaaacttg gatcgacgac 1440  
 gatgatgatg aaggaggaaa gggacatacc ctattaacat ttgctatagt gcacagatat 1500  
 cccttaataa gccaaacttat ttcacacca attttaaaat cgttagtgaa tactacatgt 1560  
 agagacaagc actttactcc cctcatgcac ctgcacca cgtctataat gtaccaatgc 1620  
 aatacactct tgtgccttat aataaatgga gctaaaccag agttcataaa caagttcaac 1680  
 gagaatgttt tgcatatagc gattgaaaaa gtttaactat gagtcacac tgaattgaga 1740  
 ggaacattat ccagcgaaca aattgaaaaa atgggtcaatg taagaagaat gatggataat 1800  
 acaacacctt taatgatcgc cttggcgagg gagaatattg tactcgctca gctttttgac 1860  
 ggtctttaca agcccaaaat aaaggtccgt ttcggttctt caaagaggct aaggatacca 1920  
 gagtttgtcc tcttaagggt cctaaaggaa tcagttgcat atttggaaac gaggaatata 1980  
 tcctacgata ttaacatcat aaaggatgca gtaatggaca acagtctttt tgaagaggag 2040  
 tacgaaatag cagcagcagg actgcgaggc aataactgcg accctgaagc agacgagaag 2100  
 actatgaaca cgtggaactt tttcaccaaa aattcaacca aatgggcaag ctctattttc 2160  
 caaaagaata ggcagaaatt tgtaaagatt gtggatggta tgaataggac atatgaagac 2220  
 tctgaatgtg caatatgctt ggatagtctg gacggggatc ttccttcagg gagaacaacg 2280  
 tgcggtcatt gcttccacaa cgtctgttgg ttatccttga taaggatgag cgggccgaat 2340  
 aatggcagcc gcgcaagagg aggaggaata aaatgcccggt cctgcagaca agtcacctgc 2400  
 ctcgaaaaaa gactaggggt tgccgactat gatattgaaa cagaggaaga acgtgacacg 2460  
 aaaaatgtcg tgccttcggt agaagaagga agaagggaat ggaggaagat ttggtgttgac 2520  
 agatatgaat ttcttgttagg tggagtgtgg acaaatgaaa taaaactata a 2571

<210> 255  
 <211> 846  
 <212> PRT  
 <213> SHRIMP

WO 01/38351

348

PCT/US00/28888

<400> 255

Met	Ala	Ser	Thr	Ser	Ser	Ser	Thr	Lys	Lys	Arg	Val	His	Glu	Glu	Asp
1				5					10					15	
Glu	Asn	Leu	Ile	Pro	Gln	Pro	Lys	Lys	Lys	Ser	Lys	Lys	Val	Leu	
			20					25					30		
Pro	Phe	Pro	Val	Asp	Lys	Tyr	Arg	Ala	Val	Asp	Lys	Lys	Val	Val	Asn
		35					40					45			
Leu	Ile	His	Lys	Ile	Leu	Asp	Gln	Glu	Lys	Asp	His	Leu	Ser	Ser	Thr
	50					55					60				
Glu	Leu	Gln	Met	Ile	Thr	Glu	Cys	Asn	Gly	Ala	Arg	Glu	Asp	Leu	Leu
65					70					75					80
Lys	His	Leu	Leu	Asp	Glu	Gly	Glu	Phe	Asn	Pro	Thr	Ile	Ile	Glu	Val
				85				90						95	
Val	Ser	Ser	Met	Pro	Ile	Glu	Thr	Ile	Tyr	Glu	Ile	Leu	Ser	Ser	Ser
			100					105					110		
Ala	Asp	Asp	Lys	Lys	Phe	Val	Gln	Ile	Ser	Leu	Ser	Met	Leu	Ile	His
		115					120					125			
Ile	Leu	Phe	Phe	Ala	Asp	Lys	Gly	Thr	Met	Trp	Val	Ser	Asn	Ala	Cys
	130					135					140				
Val	Gln	Asn	Val	Leu	Gly	Asn	Asp	Tyr	Lys	Val	Glu	Phe	Glu	Asn	Ile
145					150					155					160
Arg	Lys	Lys	Tyr	Leu	Ile	Asp	Leu	Leu	Asn	Gly	Val	Ser	Asn	His	Trp
				165					170					175	
Ser	Glu	His	Gly	Pro	Leu	Ser	His	Met	Leu	His	Ser	Ser	Ile	Pro	Ile
			180					185					190		
Val	Gln	Asp	Met	Leu	Leu	Asn	Arg	Leu	Val	Arg	Tyr	Phe	Ser	Thr	Tyr
		195					200					205			
Asp	Gly	Asp	Ala	Gln	Phe	Asp	Ile	Ser	Phe	Ile	Ile	Asn	Ser	Val	Leu
	210					215					220				
Trp	Gly	Ile	Asp	Lys	Ser	Val	Leu	Asn	Glu	Leu	Thr	Gln	Leu	Ile	Ser
225					230					235					240
Arg	Gly	Val	Phe	Ile	Val	Ser	Tyr	Val	Pro	Met	Arg	Val	Arg	Thr	Pro
				245					250					255	
Ser	Lys	Asp	Ser	Asn	Arg	Pro	Gln	Asn	Thr	Pro	Ser	Gln	Asn	Met	Ser
			260					265					270		
Ala	Leu	Gly	Met	Lys	Leu	Asn	Thr	Phe	Ser	Ser	Arg	Ile	Ser	Val	Tyr
		275					280					285			
Arg	Asn	Asn	Thr	Phe	Lys	Lys	Leu	Thr	Glu	Leu	Val	His	Asn	Phe	Asp
	290					295					300				
Tyr	Gly	Ser	Lys	Asp	Ala	Ser	Ser	Ser	Ser	Pro	Pro	Pro	Pro	Ser	Leu
305					310					315					320
Ser	Asp	Ser	Val	Asn	Thr	Phe	Val	Arg	Leu	Tyr	Thr	Asn	Tyr	Asp	Ile
				325					330					335	
Phe	Leu	Lys	Val	Ile	Ser	Asp	Trp	Lys	Met	Pro	Tyr	Gly	Phe	Phe	Lys
			340					345					350		
Lys	Thr	Phe	Asp	Val	Lys	Lys	Gly	Leu	Met	Thr	Leu	Ser	Val	Ser	Glu
		355					360					365			
Tyr	Thr	Leu	Lys	Lys	Glu	Leu	Val	Thr	Phe	Leu	Arg	Ala	Leu	Lys	Glu
	370					375					380				
Arg	Glu	Ile	Leu	Ile	Tyr	Lys	Met	Glu	Lys	Arg	Asp	Ile	Ile	Cys	Ile
385					390					395					400
Leu	Lys	Lys	Ser	Leu	Phe	Gly	Phe	Asn	Phe	Arg	Cys	Leu	Lys	Gln	Leu
				405					410					415	
Leu	Pro	Leu	Phe	Lys	His	Phe	Leu	Lys	Ile	Glu	Glu	Val	Lys	His	Ile
			420					425					430		
Ala	Arg	Phe	Val	Phe	Arg	Asp	Tyr	Ser	Leu	Met	Cys	Lys	Thr	Gln	Lys
		435					440					445			
Asp	Leu	Gln	Ser	Phe	Pro	Ala	Ile	Gln	Ser	Ala	Ser	Leu	Phe	Met	Glu
	450					455					460				
Glu	Phe	Pro	Trp	Leu	Ala	Lys	Thr	Trp	Ile	Asp	Asp	Asp	Asp	Asp	Glu
465					470					475					480



<210>	256
<211>	930
<212>	DNA
<213>	SHRIMP

<400> 256						
atggacaatc	ttatcaccaa	cgacaacatt	atcctcgtca	ctttcctgag	tggattggca	60
gtcggctgct	ccatgactat	tgggctcgca	ctggccatga	acatgctcgt	gaagtgcatt	120
gacagaacta	ctacttgcatt	ttcttgtctca	ccgtggggaga	agaataagaa	taagaagaac	180
aggaacggga	gcaaacaccga	atccagtttc	atcagccacg	tccggttcaa	cactccagat	240
aaggaccctg	acatctctga	acccatgtct	aaatctacca	cttacgatct	ggccaatgtt	300
accctccaag	tcacaaaact	ggttaacattt	tctggtccaa	cctatgctag	tcccgctaca	360
cccaggccag	ttgccaatat	acctcaacaa	caaccaacaa	gtacaataaa	aqaggaagaa	420

```

agtgtctata tgccaatgtc gagctgctcg tcgtcatttt cttctgacaa tagtcttcct 480
ctgccaacac cgccgccatc tccacctaga agcaatggcg gtgattacgt gtcatatgta 540
aacggacgac atctgaagct tccttcaaac ccaccttctc ccatcttcaa tatcaagaat 600
gaggaggagg aggatgataa tgtggaagaa catgtctacg aatacgtgcc agaagtacct 660
caacaatctc catctatcca gaagtgtatc caggaaattga aggagatgaa acacaagaaa 720
aacaccctaa ccaggagcag tagtaacaac aacaacaatg ctccacgtat aacccaagtt 780
acgtttaaga aattcccacc taacaataat aacatgtggg agaatcatgt gtatggaaac 840
actacaattg tgtcttcac accttctcct acctttattc cttcacctaa aagtatcata 900
aggaaattgt catttaagag gaaacaataa 930

```

&lt;210&gt; 257

&lt;211&gt; 305

&lt;212&gt; PRT

&lt;213&gt; SHRIMP

&lt;400&gt; 257

```

Met Asp Asn Leu Ile Thr Asn Asp Asn Ile Ile Leu Val Thr Phe Leu
1      5      10
Ser Gly Leu Ala Val Gly Cys Ser Met Thr Ile Gly Leu Ala Met Asn
20     25     30
Met Leu Val Lys Cys Ile Asp Arg Thr Thr Thr Cys Ile Ser Cys Ser
35     40     45
Pro Trp Glu Lys Asn Lys Asn Lys Lys Asn Arg Asn Gly Ser Asn Thr
50     55     60
Glu Ser Ser Phe Ile Ser His Val Arg Phe Asn Thr Pro Asp Lys Asp
65     70     75     80
Leu Asp Ile Ser Glu Pro Met Leu Lys Ser Thr Thr Tyr Asp Leu Ala
85     90     95
Asn Val Thr Pro Gln Val Thr Lys Leu Val Thr Phe Ser Gly Pro Thr
100    105    110
Tyr Asp Pro Thr Pro Arg Pro Val Ala Asn Thr Pro Gln Gln Gln Pro
115    120    125
Thr Ser Thr Asn Lys Glu Glu Glu Ser Val Tyr Met Pro Met Ser Ser
130    135    140
Cys Ser Ser Ser Phe Ser Ser Asp Asn Ser Leu Pro Leu Pro Thr Pro
145    150    155    160
Pro Pro Ser Pro Pro Arg Ser Asn Gly Gly Asp Tyr Val Ser Tyr Val
165    170    175
Asn Gly Arg His Leu Lys Leu Pro Ser Asn Pro Pro Ser Pro Ile Phe
180    185    190
Asn Ile Lys Asn Glu Glu Gly Glu Asp Asp Asn Val Glu Glu His Val
195    200    205
Tyr Glu Tyr Val Pro Glu Val Pro Gln Gln Ser Pro Ser Ile Gln Lys
210    215    220
Cys Ile Gln Glu Leu Lys Glu Met Lys His Lys Lys Asn Thr Leu Thr
225    230    235    240
Arg Ser Ser Ser Asn Asn Asn Asn Ala Pro Arg Ile Thr Gln Val
245    250    255
Thr Phe Lys Lys Phe Pro Pro Asn Asn Asn Asn Met Trp Glu Asn His
260    265    270
Val Tyr Gly Asn Thr Thr Ile Val Ser Ser Thr Pro Ser Pro Thr Phe
275    280    285
Ile Pro Ser Pro Lys Ser Ile Ile Arg Lys Leu Ser Phe Lys Arg Lys
290    295    300
Gln
305

```

&lt;210&gt; 258

&lt;211&gt; 549

&lt;212&gt; DNA

&lt;213&gt; SHRIMP

&lt;400&gt; 258

```

atgggagaaaa gtatttttcga tgctgtaagt ctggcaacaa ataatcctaa aaagtcaaac 60
tccagaaaca agaagttatt gcgggaactc aagaatatgc ggaaagattt tccttcaacc 120
tttctccagt gtcggatgat agattttcac ttttctggtg acattataga taaacattac 180
tgtcattctg taaacgtgcc agatgtttgt cctaatacaa tatttgcagt tttcttacct 240
gaagaggacc gtgccaataa ccccggtgta tacgattcta ttgaaggagt atgtataaca 300
gtcgaacaag gtgaattatg catcatcaac aagtcaagcg ttcacgagtt caatattctg 360
gtgtccttgc ataaggactt atttggtgaa gatattcttg atggaataga aactgcatca 420
agggagaagt ctcggtctat ccacctatat ctggaggctg ggcagagtat cagaacccca 480
atcccaagac cagaagggac aaatactgtg aactacacta tcgttttttc taaccaagtg 540
acgtataa                                     549

```

&lt;210&gt; 259

&lt;211&gt; 180

&lt;212&gt; PRT

&lt;213&gt; SHRIMP

&lt;400&gt; 259

```

Met Gly Glu Ser Ile Phe Asp Ala Val Ser Leu Ala Thr Asn Asn Pro
 1           5           10           15
Lys Lys Ser Asn Ser Arg Asn Lys Lys Leu Leu Arg Glu Leu Lys Asn
 20           25           30
Met Arg Lys Asp Phe Pro Ser Thr Phe Leu Gln Cys Arg Met Ile Asp
 35           40           45
Phe His Phe Ser Gly Asp Ile Ile Asp Lys His Tyr Cys His Ser Val
 50           55           60
Asn Val Pro Asp Val Val Pro Asn Thr Ile Phe Ala Val Phe Leu Pro
 65           70           75           80
Glu Glu Asp Arg Ala Asn Asn Pro Gly Asp Ser Ile Glu Gly Val Cys
 85           90           95
Ile Thr Val Glu Gln Gly Glu Leu Cys Ile Ile Asn Lys Ser Ser Val
100          105          110
His Glu Phe Asn Ile Leu Val Ser Leu His Lys Asp Leu Phe Gly Glu
115          120          125
Asp Ile Leu Asp Gly Ile Glu Thr Ala Ser Arg Glu Glu Ser Arg Ser
130          135          140
Ile His Leu Tyr Leu Glu Ala Gly Gln Ser Ile Arg Thr Pro Ile Pro
145          150          155          160
Arg Pro Glu Gly Thr Asn Thr Val Asn Tyr Thr Ile Val Phe Ser Asn
165          170          175
Gln Val Thr Val
180

```

&lt;210&gt; 260

&lt;211&gt; 3543

&lt;212&gt; DNA

&lt;213&gt; SHRIMP

&lt;400&gt; 260

```

atggaagcgg catcaaatct gcgtataact gaagggtgcgg gagtggttga tatcgacaac 60
gaggacgata tcaataataa tggtgattat agtaacttgt atgaagacga agaggatgag 120
gaggaaatga acgaggatga ggaggaggag gaggaggagg attacgaaga tgaagatgag 180
gacacgggag ttgaaatgg aagaaataaa gatcctccat cttctaaaaa acaaagtaaa 240
tttgtgagag atgtcactaa tgacatgtac gatgatgatg atgaagaaga agaggaagag 300
gaggaagagg aagatgaaga gggagaagaa ggggggtgaat atgacgaaa tctagaagac 360
gaggaagaag aaggagatga atatgaagat gacaatgaag gtgaagggga agaagatgag 420
gctgaccctg cgttattggc gctagcggca caacaagaag atgcgacaat tatacctgaa 480
aaccagtgga aaagtatagt gaacaccccc tcgccagtag ggccaaatag gcaagttctt 540

```

```

cccatgttga actttttact tgaaaatgtg aacgccatgg gcggatcagc aggtgaagaa 600
cagaagaaca aagaagatga taaccaacaa atagaacccg tagaagagga agaggatgaa 660
gaagaggaag aacaggagga ggaagaggaa gaggaagaag aacaggagga ggaagaggaa 720
gaaaaagaac ctatagaaca agaaaaaaaaat gaacccgaga aggatgaaga tgcaatagaa 780
aatgaaagcg tgcaactctca tagagttgaa tccagcccta tgagtgaagg aggtaatgat 840
gatggaatgg attattttctt ctcttcaata gcagggtggcg gcaatgataa tgaagaggat 900
gaagaagagg atgaagaaga gggagaggaa gaggaggagg aggaacctgc acaaaaaatct 960
gaagagcatg tagaaactaa agagtctgtt caatctcaca cggaatataat tgaagaagag 1020
gaagaatatg aagaatatga agatgaatct cggcacacac tcgaggatga agaaataagt 1080
actatgcate aatttaacaa tgcgccagg gttcgaagat ctccccctcc agatatacaa 1140
gagtgcgagg atgcagttgt gttcccccca atcatgaaag agacagatat tctccccag 1200
ataaaggaac cttcaccaaaa agcacctaga atgttttcta ttcttggtag tggagggtgag 1260
gaacaatatg accaactaaa cgatatagca ccaccgcctg tcccatctat cgtgactttt 1320
cctccagaca atgaaatggg agaagaaagt agagacatca tggaccaaga ttcgatgctg 1380
atgctctctc ctctctctcc acctccacca ccacctcatc aacctctca actcaaacct 1440
accaatatct ttcttctctc cctctccact ccacctacaa atcagtcctt attttcaaac 1500
aacaacaaca atccatcctt tttgagtaca gtggttgga aggttaacaa cacattaggt 1560
gggaaagagg ccgaagaaag attacacaag actatggagt ctataatctt aaaaacaagg 1620
gtaaaaactc ttctagaaac aacaaagaac cttcaatgct ctgaactggt gaaagttgtg 1680
ttccaagatc cagaaaaccc agtaaaacca tccgagaaag taatggagcg ccttaaaaaat 1740
atcatcgctg cagaattaac catgaaggca ttttttagatt ctgctgcggt gacagacatt 1800
aaatctgccg agctgttcag gaaaacaaat gagaaattgg agttgtttca gaggaagcaa 1860
atcatgtcca accctctttt ttcagcagca tacgcgtcaa cttatataat gggggaaaga 1920
gctagcaaaa taagaccatc gactcctgct ccttctctta aaaaggttga aagtatatct 1980
gaattaaacg aggatgaaac ttccatgtct tcactctgctg gtggagtatg tgctgaaggc 2040
gacgagtcta ttgctggagg aggaggagga ggaggaggag ggtggtggag 2100
cattcttcat tttatagtaa tcaaaactcaa gcaaatttgc acatggaact gataaatata 2160
ctgaaagaag acgacgataa tcaaccatgt caaacgtata agcttggtca aagactggcg 2220
ttcctaaaca acctcatatc ctttaaaaca tcaagtgcgtg tttcttggtc tagattagtc 2280
aatatgcttt cggacatagt gaccaaggca tcagtggcac tgttcggaga caccaacaaa 2340
gccaagagg attttgaaaa acatcaaact gaaacaaatg atgtatcaga tttgtctacc 2400
tcacttaaac taaaacagat gagtaagaa tctgcaaaaca taatggaaga aatggggtta 2460
ggttctatag gggcagaaat ttgtttcggg gcaatttcta caatcataga gaaacacatc 2520
aacaactat gtatggacgt ggggaagggtt acaattttct tgaacatacc aattgtacta 2580
ttaaattggc caaaagagtt cactttgtca aaagattata aagttcttct tttggatagt 2640
atttctagtt gttcttctaa aatggctgtc cctcccattt acgtcttaaa cagtatacaa 2700
tttgataagg cagtcgatga agaggatgaa gatggtaatg gaagtgaggc agagaaaaga 2760
agtgaagacg gcaacatgtt ttcagagaag gacaagaagg aagcaattcg tcgagtatac 2820
gataatataa gatacgggga cagtaatgac cgtacatcat taaaccactt tttcggtgac 2880
gcgtattctg gagtgaagta caataacagc aagaatagta tgtttgatct ccagacacaa 2940
ggcgggggaa ggtttggtgt agcatacagt gcaggctcat caatcatcga acacagatcc 3000
cctatttttg ataatgcatt aaatacacta gtcaatttca tggacaagag gaaacacttg 3060
ctaagcgcag tagtgatcaa actttttaaag aaagctaaat tgtccattga agtgactgc 3120
ataaaatata agctaaatca agcatctgag aaatacaata aaaagggtaa gcacggaaaa 3180
tctacatctg ttgtgcctat gcgaaactta atgtaccgtc cttctaaaaa tcaagatgtt 3240
tctccctcca cccagcagc agcaacagca atggacgtgc ctagcagtg atcttctcat 3300
gttggttagg aaagaacgtt ttctttttca aatgacatca attccaacat gagtagcgct 3360
agcagtggt atattgacca agaattcttc acaccttcta gacggaggac ttttatggac 3420
cttttgaata acaaactctag cgtcaattct ctagcaaagc aagtgaaaag aatgaagcac 3480
acaaaatatt acaattcttc atccaattca gaagacgacg atgaagacga ccaatacagag 3540
taa 3543

```

<210> 261  
 <211> 1174  
 <212> PRT  
 <213> SHRIMP

<400> 261  
 Met Glu Ala Asn Leu Arg Ile Thr Glu Gly Ala Gly Val Leu Asp Ile  
 1 5 10 15  
 Asp Asn Glu Asp Asp Ile Asn Asn Asn Val Asp Tyr Ser Asn Leu Tyr

			20							25							30				
Glu	Asp	Glu	Glu	Asp	Glu	Glu	Glu	Met	Asn	Glu	Asp	Glu	Glu	Glu	Glu	Glu					
		35					40						45								
Glu	Glu	Glu	Asp	Tyr	Glu	Asp	Glu	Asp	Glu	Asp	Thr	Gly	Val	Arg	Asn						
	50					55					60										
Gly	Arg	Asn	Lys	Asp	Pro	Pro	Ser	Ser	Lys	Lys	Gln	Ser	Lys	Phe	Val						
65					70				75					80							
Arg	Asp	Val	Thr	Asn	Asp	Met	Tyr	Asp	Asp	Asp	Asp	Glu	Glu	Glu	Glu						
				85					90				95								
Glu	Glu	Glu	Glu	Glu	Glu	Asp	Glu	Glu	Gly	Glu	Glu	Gly	Gly	Glu	Tyr						
			100					105					110								
Asp	Gly	Asn	Leu	Glu	Asp	Glu	Glu	Glu	Glu	Gly	Asp	Glu	Tyr	Glu	Asp						
		115					120					125									
Asp	Asn	Glu	Gly	Glu	Gly	Glu	Glu	Asp	Glu	Ala	Asp	Pro	Ala	Leu	Leu						
	130					135					140										
Ala	Ala	Gln	Gln	Glu	Asp	Ala	Thr	Ile	Ile	Pro	Glu	Asn	Gln	Trp	Lys						
145					150				155					160							
Ser	Ile	Val	Asn	Thr	Pro	Ser	Pro	Val	Gly	Pro	Asn	Arg	Gln	Val	Leu						
			165					170					175								
Pro	Met	Leu	Asn	Phe	Leu	Leu	Glu	Asn	Val	Asn	Ala	Met	Gly	Gly	Ser						
			180					185					190								
Ala	Gly	Glu	Glu	Gln	Lys	Asn	Lys	Glu	Asp	Asp	Asn	Gln	Gln	Ile	Glu						
		195					200					205									
Pro	Val	Glu	Glu	Glu	Glu	Asp	Glu	Glu	Glu	Glu	Glu	Gln	Glu	Glu	Glu						
	210					215					220										
Glu	Glu	Glu	Glu	Glu	Glu	Gln	Glu	Glu	Glu	Glu	Glu	Glu	Lys	Glu	Pro						
225					230					235				240							
Ile	Glu	Gln	Glu	Lys	Asn	Glu	Pro	Glu	Lys	Asp	Glu	Asp	Ala	Ile	Glu						
			245						250				255								
Asn	Glu	Ser	Val	His	Ser	His	Arg	Val	Glu	Ser	Ser	Pro	Met	Ser	Glu						
			260					265					270								
Gly	Gly	Asn	Asp	Asp	Gly	Met	Asp	Tyr	Phe	Phe	Ser	Ser	Ile	Ala	Gly						
		275					280				285										
Gly	Gly	Asn	Asp	Asn	Glu	Glu	Asp	Glu	Glu	Glu	Asp	Glu	Glu	Glu	Gly						
	290					295					300										
Glu	Glu	Glu	Glu	Glu	Glu	Glu	Pro	Ala	Gln	Lys	Ser	Glu	Glu	His	Val						
305						310				315				320							
Glu	Thr	Lys	Glu	Ser	Val	Gln	Ser	His	Thr	Glu	Tyr	Ile	Glu	Glu	Glu						
			325						330				335								
Glu	Glu	Tyr	Glu	Glu	Tyr	Glu	Asp	Glu	Ser	Arg	His	Thr	Leu	Glu	Asp						
			340					345					350								
Glu	Glu	Ile	Ser	Thr	Met	His	Gln	Phe	Asn	Asn	Ala	Pro	Arg	Val	Arg						
		355					360				365										

Asn Thr Leu Gly Gly Lys Glu Ala Glu Glu Arg Leu His Lys Thr Met  
 515 520 525  
 Glu Ser Ile Ile Leu Lys Thr Arg Val Lys Thr Leu Leu Glu Thr Thr  
 530 535 540  
 Lys Asn Leu Gln Cys Ser Glu Leu Val Lys Val Val Phe Gln Asp Pro  
 545 550 555 560  
 Glu Asn Pro Val Lys Pro Ser Glu Lys Val Met Glu Arg Leu Lys Asn  
 565 570 575  
 Ile Ile Ala Ala Glu Leu Thr Met Lys Ala Phe Leu Asp Ser Ala Ala  
 580 585 590  
 Val Thr Asp Ile Lys Ser Ala Glu Leu Phe Arg Lys Thr Asn Glu Lys  
 595 600 605  
 Leu Glu Leu Phe Gln Arg Lys Gln Ile Met Ser Asn Pro Leu Phe Ser  
 610 615 620  
 Ala Ala Tyr Ala Ser Thr Tyr Ile Met Gly Glu Arg Ala Ser Lys Ile  
 625 630 635 640  
 Arg Pro Ser Thr Pro Ala Pro Ser Leu Lys Lys Val Glu Ser Ile Ser  
 645 650 655  
 Glu Leu Asn Glu Asp Glu Thr Ser Met Ser Ser Ser Ala Gly Gly Val  
 660 665 670  
 Cys Ala Glu Gly Asp Glu Ser Ile Ala Gly Gly Gly Gly Gly Gly  
 675 680 685  
 Gly Gly Gly Gly Glu Val Val Glu His Ser Ser Phe Tyr Ser Asn Gln  
 690 695 700  
 Thr Gln Ala Asn Leu His Met Glu Leu Ile Asn Ile Leu Lys Glu Asp  
 705 710 715 720  
 Asp Asp Asn Gln Pro Cys Gln Thr Tyr Lys Leu Gly Gln Arg Leu Ala  
 725 730 735  
 Phe Leu Asn Asn Leu Ile Ser Phe Lys Thr Ser Ser Ala Val Ser Trp  
 740 745 750  
 Ser Arg Leu Val Asn Met Leu Ser Asp Ile Val Thr Lys Ala Ser Val  
 755 760 765  
 Phe Gly Asp Thr Asn Lys Ala Gln Glu Asp Phe Glu Lys His Gln Thr  
 770 775 780  
 Glu Thr Asn Asp Val Ser Asp Leu Ser Thr Ser Ser Lys Leu Lys Gln  
 785 790 795 800  
 Met Ser Lys Glu Ser Ala Asn Ile Met Glu Glu Met Gly Leu Gly Ser  
 805 810 815  
 Ile Gly Ala Glu Ile Cys Phe Gly Ala Ile Ser Thr Ile Ile Glu Lys  
 820 825 830  
 His Ile Asn Lys Leu Cys Met Asp Val Gly Arg Leu Thr Ile Phe Leu  
 835 840 845  
 Asn Ile Pro Ile Val Leu Leu Asn Trp Pro Lys Glu Phe Thr Leu Ser  
 850 855 860  
 Lys Asp Tyr Lys Val Leu Leu Leu Asp Ser Ile Ser Ser Cys Ser Ser  
 865 870 875 880  
 Lys Met Ala Val Pro Ile Tyr Val Leu Asn Ser Ile Gln Phe Asp  
 885 890 895  
 Lys Ala Val Asp Glu Glu Asp Glu Asp Gly Asn Gly Ser Glu Ala Glu  
 900 905 910  
 Lys Arg Ser Glu Asp Gly Asn Met Phe Ser Glu Lys Asp Lys Lys Glu  
 915 920 925  
 Ala Ile Arg Arg Val Tyr Asp Asn Ile Arg Tyr Gly Asp Ser Asn Asp  
 930 935 940  
 Arg Thr Ser Leu Asn His Phe Phe Gly Asp Ala Tyr Ser Gly Val Ser  
 945 950 955 960  
 Asn Asn Asn Ser Lys Asn Ser Met Phe Asp Leu Gln Thr Gln Gly Gly  
 965 970 975  
 Gly Arg Phe Gly Val Ala Tyr Ser Ala Gly Ser Ser Ile Ile Glu His  
 980 985 990  
 Arg Ser Pro Ile Phe Asp Asn Ala Leu Asn Thr Leu Val Asn Phe Met

```

          995              1000              1005
Asp Lys Arg Lys His Leu Leu Ser Ala Val Val Ile Lys Leu Leu Lys
    1010              1015              1020
Lys Ala Lys Leu Ser Ile Glu Val Tyr Cys Ile Lys Tyr Lys Leu Asn
1025              1030              1035              1040
Gln Ala Ser Glu Lys Tyr Asn Lys Lys Gly Lys His Gly Lys Ser Thr
    1045              1050              1055
Ser Val Val Pro Met Arg Asn Leu Met Tyr Arg Pro Ser Lys Asn Gln
    1060              1065              1070
Asp Val Ser Pro Ser Thr Pro Ala Ala Thr Ala Met Asp Val Pro
    1075              1080              1085
Ser Ser Val Ser Ser His Val Gly Arg Lys Arg Thr Phe Ser Phe Ser
    1090              1095              1100
Asn Asp Ile Asn Ser Asn Met Ser Ser Ala Ser Ser Val Tyr Ile Asp
1105              1110              1115              1120
Gln Glu Ser Ser Thr Pro Ser Arg Arg Arg Thr Phe Met Asp Leu Leu
    1125              1130              1135
Asn Asn Lys Ser Ser Val Asn Ser Leu Ala Lys Gln Val Lys Arg Met
    1140              1145              1150
Lys His Thr Lys Tyr Tyr Asn Ser Ser Ser Asn Ser Glu Asp Asp Asp
    1155              1160              1165
Glu Asp Asp Gln Tyr Glu
    1170

```

<210> 262  
 <211> 786  
 <212> DNA  
 <213> SHRIMP

```

<400> 262
atgtctaattg ggcgaactat aagcgatgag aggttgatcc tcatccttga taaaattgta 60
gaaagacgtg gtgtttctaa tctatctgaa ttgttgatac accccataac caaacacata 120
aacgaattgt tgaagaacac tgtaagacat ggagacagag tttacatgaa ggatgcagaa 180
ctggatgtga gatctgcgct agaagacata aaaaaggatt gtgtttttaa ggcaattgaa 240
aaacaaggaa tagatgttag acaaataata actgattact tggctaaacg aaaactaacg 300
caaaatcttg tacattggta tcggcccccata atatcttgca cagatataga cgaaaaaatt 360
caacaagaaa ctggtcaagt agggcggtgt agtgttgcta cgtacaattt gagaattggt 420
ggtgacgatg gagaatttac aaggtacgat ttctccattc ccttgggaga ttttaaaata 480
acggcaaaat tgttcggttc cataaatgat gaggatgtag atgcagtgat tcttgtgtct 540
cgtagtgcg tagttaatga cgtgctaagc tttagaat ttaatcgaac aggagaacgc 600
gtagtcatat tctttaatgt gattgttgaa gggaagagta aagatatgta tattgtatgt 660
aaatctagat ataacacac ccatatacta aacggagaat ctgcaacata cgctgttaaa 720
cgtataaaaa gaggcgatac aagggacgat atattgtttg caatcactgc ttttaaggag 780
gagtaa 786

```

<210> 263  
 <211> 261  
 <212> PRT  
 <213> SHRIMP

```

<400> 263
Met Ser Asn Gly Ala Thr Ile Ser Asp Glu Arg Leu Ile Leu Ile Leu
  1          5          10          15
Asp Lys Ile Val Glu Arg Arg Gly Val Ser Asn Leu Ser Glu Leu Leu
  20          25          30
Ile His Pro Ile Thr Lys His Ile Asn Glu Leu Leu Lys Asn Thr Val
  35          40          45
Arg His Gly Asp Arg Val Tyr Met Lys Asp Ala Glu Leu Asp Val Arg
  50          55          60
Ser Arg Leu Glu Asp Ile Lys Lys Asp Cys Val Leu Lys Ala Ile Glu

```

<400> 265															
Met	Ser	Ser	Asn	Gly	Asp	Glu	Pro	Ala	Val	Thr	Glu	Ala	Glu	Ile	Ala
1				5					10					15	
Ser	Val	Glu	Ala	Gln	Leu	Gly	Ala	Ala	His	His	Asp	Asn	Ser	Trp	Ile
			20					25					30		
Thr	Arg	Lys	Ser	Asp	Gln	Leu	Lys	Tyr	Arg	Leu	Gly	Ala	Ile	Ala	Tyr
		35					40					45			



Ser Val Ala Lys Asn Ala Ser Ile Lys Tyr Ile Glu Asp Gln Val Arg  
 50 55 60  
 Gln Glu Ile Asn Ser His Leu Thr Asn Val Met Thr Phe Glu His Leu  
 65 70 75 80  
 Tyr Glu Asp Ala Phe Asn Pro Val Ile Cys Glu Ala Ile Phe Glu Lys  
 85 90 95  
 Gly Ile Pro Val Val Met Glu Lys Val Tyr Asp Val Asn Arg Arg Ile  
 100 105 110  
 Met Glu Pro Arg Glu Asp Phe Ile Thr Glu Ile Leu Lys Glu Glu Arg  
 115 120 125  
 Trp Arg Arg Tyr Ile Pro Gly Phe Tyr His Thr Ser Phe Ser Phe Lys  
 130 135 140  
 Tyr Asn Thr Ile Ala Phe Thr Asp Ser Ser Thr Ser Phe Ser Val Pro  
 145 150 155 160  
 Ile Asn Asp Lys His Met Leu Ser Ile Thr Pro Pro Gly Ala Ala Gln  
 165 170 175  
 Gly Asp Leu Ile Asp Leu Ser Leu Ser Phe Lys Ile Asp Ser Ser Ala  
 180 185 190  
 Lys Thr Leu Thr Leu Glu Phe Asn Arg Lys Ser Thr Phe Ala Gly Ile  
 195 200 205  
 Val Asn Arg Pro Lys Ser Val Val Ile Leu Ser Asn Leu Arg Asn Ser  
 210 215 220  
 Asp Ser Ser Asp Asn Ile Gly Asp Tyr Leu Lys Arg Asn Asp Pro Ile  
 225 230 235 240  
 Tyr Ile Ser His Asp Thr Asn Gly Ile Ile Asn Pro Ser Glu Asp Ser  
 245 250 255  
 Ala Ser Leu Ile Thr Ile His Met Pro Glu Ile Glu Asn Ala Ser Asp  
 260 265 270  
 Asp Leu Tyr Ile Asp Phe Asn Leu Phe Val Phe  
 275 280

<210> 266  
 <211> 1302  
 <212> DNA  
 <213> SHRIMP

<400> 266  
 atggcacttt caaacaatgg aggaatatac attgtttttg cggttattgt tttggtaata 60  
 ggagcttcta ttgccctctt ctttgctatc tcgggcgtag ggaagggaac tctacattca 120  
 aatgccaaaa caaaaaagag taagaaatat aaattagact ctaaatacac tgacgatgat 180  
 gaaaaaactg acgacgataa taataataat ggaggaggag ggggaggaac agttgatgtt 240  
 atcaatgaga cagcgcttca acgtcaaacg agagagcatt ttgcaagaac tcttgaaaaa 300  
 gctgaggatg aattcttcac caaattagca gatcaggaat ttgacacata caaatcagaa 360  
 aacgtatggt taataaagga taaaataaca gatggaaaag tttcaatccc tgaaggtgac 420  
 ataaacgtcc ccgatgtcgg acaggcaatt gctgatgaaa acttgttcga tctcataggg 480  
 acgaaccatg acgaagtcaa ggaaacgatg gatgaagttg ttgcacaaaa atctaccaat 540  
 atcacttacg aacaactcgt aatagacttg accaatattt tattgtttgg tacagtaaca 600  
 gttgatcctt ctgatgaaaa tggggatgaa agcctacaga gatcaacaga cccagacgca 660  
 gaaatggtga tgttgacaac aacaccttct tcacaactag ctagacaaca acaacctcct 720  
 caacctacac ctgattacct tgcccggtag tcaaaggaat tgggtgataaa taatatacga 780  
 ggagggttta tcagtgatcg tgatatgcgc acttggcaag gacgaatgtc tgtacatgtc 840  
 aacatgaaac agaggacatt taatgttatt agtgcagcaa cgaatctgga ttctctacaa 900  
 gttggattag aacccgtgct acaaaaacaa ggtagagcag ctgtgggagg acgtattgaa 960  
 aaagcccggg tagagttttc atttgtagta gaaggtaacc gtgtacgggt atacgctaca 1020  
 aacaaaacag aggactgttt ttgtagttta ctccccaact gttataatgt taaaaaggca 1080  
 tcagactatt ggataagctc tgcaagcaca gctaaggaaa aaacgtactt gttttattgct 1140  
 aataaaaaatg atgaacaag tttcttctat aactttgagg aaggtgttga agaaattgac 1200  
 ctggacattt ttatgacaat agattgtgca cctaactctt ctttcattaa aaatttacca 1260  
 agacctatta cagataataa tataatggtt gcactgtcat aa 1302

WO 01/38351

358

PCT/US00/28888

<210> 267  
<211> 431  
<212> PRT  
<213> SHRIMP

<400> 267

Met	Ala	Leu	Ser	Asn	Asn	Gly	Gly	Ile	Tyr	Ile	Val	Phe	Ala	Val	Ile
1				5				10						15	
Val	Leu	Val	Ile	Gly	Ala	Ser	Ile	Ala	Leu	Phe	Phe	Ala	Ile	Ser	Gly
			20					25					30		
Val	Gly	Lys	Gly	Thr	Leu	His	Ser	Asn	Ala	Lys	Thr	Lys	Lys	Ser	Lys
		35					40					45			
Lys	Tyr	Lys	Leu	Asp	Ser	Lys	Tyr	Thr	Asp	Asp	Asp	Glu	Lys	Thr	Asp
	50					55				60					
Asp	Asp	Asn	Asn	Asn	Asn	Gly	Gly	Gly	Gly	Gly	Thr	Val	Asp	Val	
65					70				75					80	
Ile	Asn	Glu	Thr	Ala	Leu	Gln	Arg	Gln	Thr	Arg	Glu	His	Phe	Ala	Arg
				85				90						95	
Thr	Leu	Glu	Lys	Ala	Glu	Asp	Glu	Phe	Phe	Thr	Lys	Leu	Ala	Asp	Gln
			100					105					110		
Glu	Phe	Asp	Thr	Tyr	Lys	Ser	Glu	Asn	Val	Trp	Leu	Ile	Lys	Asp	Lys
		115					120					125			
Ile	Thr	Asp	Gly	Lys	Val	Ser	Ile	Pro	Glu	Gly	Asp	Ile	Asn	Val	Pro
	130					135					140				
Asp	Val	Gly	Gln	Ala	Ile	Ala	Asp	Glu	Asn	Leu	Phe	Asp	Leu	Ile	Gly
145					150					155					160
Thr	Asn	His	Asp	Glu	Val	Lys	Glu	Thr	Met	Asp	Glu	Val	Val	Ala	Gln
				165				170						175	
Lys	Ser	Thr	Asn	Ile	Thr	Tyr	Glu	Gln	Leu	Val	Ile	Asp	Leu	Thr	Asn
			180					185					190		
Ile	Leu	Leu	Phe	Gly	Thr	Val	Thr	Val	Asp	Pro	Ser	Asp	Glu	Asn	Gly
		195					200					205			
Asp	Glu	Ser	Leu	Gln	Arg	Ser	Thr	Asp	Pro	Asp	Ala	Glu	Met	Val	Met
	210					215				220					
Leu	Thr	Thr	Thr	Pro	Ser	Gln	Leu	Ala	Arg	Gln	Gln	Gln	Pro	Pro	
225					230					235				240	
Gln	Pro	Thr	Pro	Asp	Tyr	Leu	Ala	Arg	Tyr	Ser	Lys	Glu	Leu	Val	Ile
				245					250					255	
Asn	Asn	Ile	Arg	Gly	Gly	Phe	Ile	Ser	Asp	Arg	Asp	Met	Arg	Thr	Trp
			260					265					270		
Gln	Gly	Arg	Met	Ser	Val	His	Val	Asn	Met	Lys	Gln	Arg	Thr	Phe	Asn
		275					280					285			
Val	Ile	Ser	Ala	Ala	Thr	Asn	Leu	Asp	Ser	Leu	Gln	Val	Gly	Leu	Glu
	290					295					300				
Pro	Val	Leu	Gln	Lys	Gln	Gly	Arg	Ala	Ala	Val	Gly	Gly	Arg	Ile	Glu
305					310					315				320	
Lys	Ala	Arg	Ile	Glu	Phe	Ser	Phe	Val	Val	Glu	Gly	Asn	Arg	Val	Arg
				325				330						335	
Val	Tyr	Ala	Thr	Asn	Lys	Thr	Glu	Asp	Cys	Phe	Cys	Ser	Leu	Leu	Pro
			340				345						350		
Asn	Cys	Tyr	Asn	Val	Lys	Lys	Ala	Ser	Asp	Tyr	Trp	Ile	Ser	Ser	Ala
		355					360					365			
Ser	Thr	Ala	Lys	Glu	Lys	Thr	Tyr	Leu	Phe	Ile	Ala	Asn	Lys	Asn	Asp
	370					375					380				
Glu	Thr	Ser	Phe	Phe	Tyr	Asn	Phe	Glu	Glu	Gly	Val	Glu	Glu	Ile	Asp
385					390					395				400	
Leu	Asp	Ile	Phe	Met	Thr	Ile	Asp	Cys	Ala	Pro	Asn	Leu	Pro	Phe	Ile
				405					410					415	
Lys	Asn	Leu	Pro	Arg	Pro	Ile	Thr	Asp	Asn	Asn	Ile	Met	Val	Ser	
			420					425					430		

WO 01/38351

PCT/US00/28888

359

<210> 268  
<211> 207  
<212> DNA  
<213> SHRIMP

<400> 268  
atgtctgata tgaccagaaa catcatcgtc ggcttggccg ttgtcgtcat cgcattgagc 60  
atggtcgctt tcatgctttc tgttactcct gcacttaccg gattcctcct aggtttgggt 120  
gtatcagcac taggagttac actcttttga tgtcccacta tgaaatctcc aggggggagga 180  
aatgctacaa tcaaccccggt ggcataa 207

<210> 269  
<211> 68  
<212> PRT  
<213> SHRIMP

<400> 269  
Met Ser Asp Met Thr Arg Asn Ile Ile Val Gly Leu Ala Val Val Val  
1 5 10 15  
Ile Ala Leu Ser Met Val Ala Phe Met Leu Ser Val Thr Pro Ala Leu  
20 25 30  
Thr Gly Phe Leu Leu Gly Leu Gly Val Ser Ala Leu Gly Val Thr Leu  
35 40 45  
Phe Gly Cys Pro Thr Met Lys Ser Pro Gly Gly Gly Asn Ala Thr Ile  
50 55 60  
Asn Pro Val Ala  
65

<210> 270  
<211> 552  
<212> DNA  
<213> SHRIMP

<400> 270  
atgttccaga aatgggtttga atcgttttctg gattcttccc gacctagata tctggatacg 60  
acatgtgtat gctcagttta ttcataatttt tccccttgtc ggaaacatat aaaattttcc 120  
acatcgcatt cgcagtgagg tataaaaaatc catcctccat caatattgaa ccataataact 180  
tcctctccca ccagtggaaa gatgtgtaac caccaccaca agagattgta cctgagcact 240  
gacgaccata cgagatggta tgacaaaaat acatcatgca tctatcttga agatattgga 300  
ggagtacaat tcatgggtata cgagttccat ctaacaccaa agaacaatca actattctcc 360  
ttccctgttc acctccaaat acacaacagg aatactgaga aaacatccct cctcgtattt 420  
gaaaatgaag aagatatgag ggtcaggaac attcatccaa aatccaagat attgatcccc 480  
gtgtccaaag acacagtgtc ttagagagaat gggtttcggt acaagggtgaa aattgtatta 540  
tcaaacaaat aa 552

<210> 271  
<211> 183  
<212> PRT  
<213> SHRIMP

<400> 271  
Met Phe Gln Lys Trp Phe Glu Ser Phe Leu Asp Ser Ser Arg Pro Arg  
1 5 10 15  
Tyr Leu Asp Thr Thr Cys Val Cys Ser Val Tyr Ser Tyr Phe Ser Pro  
20 25 30  
Cys Arg Lys His Ile Lys Phe Ser Thr Ser His Ser His Glu Gly Ile  
35 40 45  
Lys Ile His Pro Pro Ser Ile Leu Asn His Asn Thr Ser Ser Pro Thr  
50 55 60

WO 01/38351

360

PCT/US00/28888

Ser Gly Lys Met Cys Asn His His His Lys Arg Leu Tyr Leu Ser Thr  
65 70 75 80  
Asp Asp His Thr Arg Trp Tyr Asp Lys Asn Thr Ser Cys Ile Tyr Leu  
85 90 95  
Glu Asp Ile Gly Gly Val Gln Phe Met Val Tyr Glu Phe His Leu Thr  
100 105 110  
Pro Lys Asn Asn Gln Leu Phe Ser Phe Pro Val His Leu Gln Ile His  
115 120 125  
Asn Arg Asn Thr Glu Lys Thr Ser Leu Leu Val Phe Glu Asn Glu Glu  
130 135 140  
Asp Met Arg Val Arg Asn Ile His Pro Lys Ser Lys Ile Leu Ile Pro  
145 150 155 160  
Val Ser Lys Asp Thr Val Leu Val Glu Asn Gly Phe Arg Tyr Lys Val  
165 170 175  
Lys Ile Val Leu Ser Asn Lys  
180

<210> 272  
<211> 684  
<212> DNA  
<213> SHRIMP

<400> 272  
atggactcac ttataagcaa attggaaaac atattctcca ttgccgagca ggactttttc 60  
aacgcggaca gcatgttcat gcaaaccatg ctctcccta cgcacgccat gttcaccgat 120  
tgcgagtctc cattgtacaa gaacaagtcg ggagggaaga atattgtcac cgatgttgga 180  
gagagtgtac tgccttcttc ttcggacgaa aagatgagct tcaaagtgcg gtcccacgta 240  
ctcaggcgat tccctgtcct acttcattgc aactacaagc agacgaatac gcccctgtgg 300  
aaggagcttt acaagcacgg gaagtttgcc ctccctcgcg acctggtggt attctccaac 360  
ccattccacc ccaatatccc cgccatgccg ttgtataaat ccccatttg tgacaccact 420  
ggaaaatcta tcattatgag tgaagtcatg accaaggagc ttttgtacaa gttggccgac 480  
aaagatatgg gccaatcttt tgctgtattg aatgttaacta accccattac tggagattct 540  
ttcctccatt actttgcagg aggaaatacc atgagggatg gggaagggga taaaatctgc 600  
acatctgctg atgtgttacg cattattgct gagataacaa tacagaaaac tggcaagatg 660  
ccatatgaat tgatgaagaa ataa 684

<210> 273  
<211> 227  
<212> PRT  
<213> SHRIMP

<400> 273  
Met Asp Ser Leu Ile Ser Lys Leu Glu Asn Ile Phe Ser Ile Ala Glu  
1 5 10 15  
Gln Asp Phe Phe Asn Ala Asp Ser Met Phe Met Gln Thr Met Leu Leu  
20 25 30  
Pro Thr Asp Ala Met Phe Thr Asp Cys Glu Ser Pro Leu Tyr Lys Asn  
35 40 45  
Lys Ser Gly Gly Lys Asn Ile Val Thr Asp Val Gly Glu Ser Val Leu  
50 55 60  
Ser Ser Ser Ser Asp Glu Lys Met Ser Phe Lys Val Leu Ser His Val  
65 70 75 80  
Leu Arg Arg Phe Pro Val Leu Leu His Cys Asn Tyr Lys Gln Thr Asn  
85 90 95  
Thr Pro Leu Trp Lys Glu Leu Tyr Lys His Gly Lys Phe Ala Leu Leu  
100 105 110  
Gly Asp Leu Val Leu Phe Ser Asn Pro Phe His Pro Asn Ile Pro Ala  
115 120 125  
Met Pro Phe Asp Lys Ser Pro Ile Cys Asp Thr Thr Gly Lys Ser Ile  
130 135 140

Ile Met Ser Glu Val Met Thr Lys Glu Leu Leu Tyr Lys Leu Ala Asp  
 145 150 155 160  
 Lys Asp Ile Gly Gln Phe Phe Ala Val Leu Asn Val Thr Asn Pro Ile  
 165 170 175  
 Thr Gly Asp Ser Phe Leu His Tyr Phe Ala Gly Gly Asn Thr Met Arg  
 180 185 190  
 Asp Gly Glu Gly Asp Lys Ile Cys Thr Ser Ala Asp Val Leu Arg Ile  
 195 200 205  
 Ile Ala Glu Ile Thr Ile Gln Lys Thr Gly Lys Met Pro Tyr Glu Leu  
 210 215 220  
 Met Lys Lys  
 225

<210> 274  
 <211> 2193  
 <212> DNA  
 <213> SHRIMP

<400> 274  
 atggagggtg gggaccaacg gacaaaactt acgccagcaa ccgtgatggg actttaccaa 60  
 tcgaaaacgc caggagaagg agaaggagga gaaggaggag ggcaattcaa gataccttca 120  
 gccatagctg tgaaatcttg ttgctctaaa aacgctactc gccgatcccc tccctcagat 180  
 tctccttatt ctcttaggcc catgaagaga ctaaaagaaga ataatggaga ggtgggagga 240  
 aaagcaccgc ctctgtgtgac tttgaggctc cgcgaggact acgagagcac accttacaac 300  
 tttaatagaa ataagaagaa gaggcctatt actattgatg aaaatcaatt tgcaacatta 360  
 aatccaacgt atgcgacaga cattatcaag aagcagcaat tgccttctgt tagtgccgcg 420  
 tctgtgttga ggaagcaccg cgccaatgcc gacaccagat acagaaaaag attctctcat 480  
 ccaaattgtg caaaattctc tactgtcaat ttgaaggcta gagactatac tccactgtct 540  
 gtctctcggt cccatgtcaa ggggccaaaa cacttgaaat cttcttgtga taccgtgact 600  
 gaaacaaatg tagtaaaagag gaacttttct tccattgaca agtgggtcaa gctagaaaaa 660  
 cccccgtgtt actttgcagt ggagagggtc gataccaata ttgcagccgg tctagaatct 720  
 ccgttccatt tgattagaca ggccgcaaaa ttaggcctca tttctgacgt gcaagatgtg 780  
 tcgtccaact acgagaccat aaaacagagc tgtattgacg caaaggaaaa agcgtccaag 840  
 tttttgtggt ctaacaaccg tactaaaaca ccccttccat cttggtggcc tgttgggttt 900  
 ggtagtaaaa acctatccgt tttagacact agccctctct tgaactggaa caggttatgc 960  
 aagaataatg gtaaaagggtg gataaaaacc atgagcatcg atcacatggc aaagaatggt 1020  
 ttttaagctt cccctggagc atgtgaatct atattggaga agaaaactac actcttgggg 1080  
 gaggtcactg cccaatgtaa gaaatgggaa agttaccgca gaaatattcc tgtaccagca 1140  
 cacgtccaac cagaatatgc ttctcaagtc gtaatgattg gaccatctga attatatctc 1200  
 gaagttaaag tcggggtata ttacatgctt gaaactggaa aagttatcaa gtttatgacg 1260  
 gacaaggaaa tgtactgtga atttgtattt gaaactgttt ttagtcacgc tcttgaggga 1320  
 agaatgaaag gcgcagtagg tgtgagaaag atgtgtgttg aagggtttttg tgtcgagatg 1380  
 gattttgcag gcatttctgt gattgatgta ttaaattggag acctgaaatg taaaatggac 1440  
 gagaatgttg tacagcaacc taaccctctg actacttcc ccaagccagc cgctgagctc 1500  
 atgcaagatc atggcagctt gtgtaggatg agggatactc tgtacgggtg taggatgctt 1560  
 caagctactg gccgcctgcc tgaaggctca caatctaaat gcaagaaacc cattacggat 1620  
 tcaatttcag ccatagctat cggttgaaaa atgagggaga gaatgttaaa ccaattgccc 1680  
 tttgttttgg tagaaattgt aaatattgtc actcggttgt ctcaacaagg attagtgaat 1740  
 ccggacataa aaagtgacaa tatagtaatt gatggaataa ctggtcaacc taagatgatt 1800  
 gattttgggt taattgtacc atgtaaaaag tactacaatt ttaaatgttg gggaaactgat 1860  
 gagaggttct ttagtaacca tcctcataca gctcctgaat ttattaacag tgagtgtgtg 1920  
 tcagaaactg ccatgacttt tgggttggtc tatttgttaa tagacatgtt gtccattttg 1980  
 attaagagaa ctgcagattt gtctgccaat tctatctata caaacattcc atttttgtct 2040  
 attgtatcta aaatgtatga ccaggaaaag accaataggc cgagagcgta tgaaattgcg 2100  
 cctgtaattg gtgcatgttt ccggttcaag gataatattg ctaaaccttt ccagtcacct 2160  
 aaacattcat tgtatagcaa gaaggttaag tag 2193

<210> 275  
 <211> 724  
 <212> PRT

WO 01/38351

PCT/US00/28888

362

<213> SHRIMP

<400> 275

Met	Glu	Gly	Gly	Asp	Gln	Arg	Thr	Lys	Leu	Thr	Pro	Ala	Thr	Val	Met
1				5					10					15	
Gly	Gln	Ser	Lys	Thr	Pro	Gly	Glu	Gly	Glu	Gly	Gly	Glu	Gly	Gly	Gly
			20					25					30		
Gln	Phe	Lys	Ile	Pro	Ser	Ala	Ile	Ala	Val	Lys	Ser	Cys	Cys	Ser	Lys
		35					40					45			
Asn	Ala	Thr	Arg	Arg	Ser	Pro	Pro	Ser	Asp	Ser	Pro	Tyr	Ser	Leu	Arg
	50					55					60				
Pro	Met	Lys	Arg	Leu	Lys	Lys	Asn	Asn	Gly	Glu	Val	Gly	Gly	Lys	Ala
	65				70				75						80
Pro	Pro	Pro	Val	Thr	Leu	Arg	Leu	Arg	Glu	Asp	Tyr	Glu	Ser	Thr	Pro
				85					90					95	
Tyr	Asn	Phe	Asn	Arg	Asn	Lys	Lys	Lys	Arg	Pro	Ile	Thr	Ile	Asp	Glu
			100					105					110		
Asn	Gln	Phe	Ala	Thr	Leu	Asn	Pro	Thr	Tyr	Ala	Thr	Asp	Ile	Ile	Lys
		115					120					125			
Lys	Gln	Gln	Leu	Pro	Ser	Val	Ser	Ala	Ala	Ser	Val	Leu	Arg	Lys	His
	130					135					140				
Arg	Ala	Asn	Ala	Asp	Thr	Gln	Tyr	Arg	Lys	Arg	Phe	Ser	His	Pro	Asn
	145					150				155					160
Cys	Ala	Lys	Phe	Ser	Thr	Val	Asn	Leu	Lys	Ala	Arg	Asp	Tyr	Thr	Pro
				165					170					175	
Leu	Ser	Val	Leu	Arg	Ser	His	Val	Lys	Gly	Pro	Lys	His	Leu	Lys	Ser
			180					185					190		
Ser	Cys	Asp	Thr	Val	Thr	Glu	Thr	Asn	Val	Val	Lys	Arg	Asn	Phe	Ser
		195					200					205			
Ser	Ile	Asp	Lys	Trp	Val	Lys	Leu	Glu	Lys	Pro	Pro	Cys	Tyr	Phe	Ala
	210					215					220				
Val	Ala	Glu	Ala	Asp	Thr	Asn	Ile	Ala	Ala	Gly	Leu	Glu	Ser	Pro	Phe
	225				230					235					240
His	Leu	Ile	Arg	Gln	Ala	Ala	Lys	Leu	Gly	Leu	Ile	Ser	Asp	Val	Gln
				245					250					255	
Asp	Val	Ser	Ser	Asn	Tyr	Glu	Thr	Ile	Lys	Gln	Ser	Cys	Ile	Asp	Ala
			260					265					270		
Lys	Glu	Lys	Ala	Ser	Lys	Phe	Leu	Trp	Ser	Asn	Asn	Arg	Thr	Lys	Gln
		275					280					285			
Pro	Pro	Ser	Ser	Trp	Trp	Pro	Val	Gly	Phe	Gly	Ser	Lys	Asn	Leu	Ser
	290					295					300				
Val	Leu	Asp	Thr	Ser	Pro	Leu	Leu	Asn	Trp	Asn	Arg	Leu	Cys	Lys	Asn
	305				310					315					320
Asn	Gly	Lys	Gly	Trp	Ile	Lys	Thr	Met	Ser	Ile	Asp	His	Met	Ala	Lys
			325						330					335	
Asn	Val	Phe	Lys	Leu	Ser	Pro	Gly	Ala	Cys	Glu	Ser	Ile	Lys	Lys	Thr
			340					345					350		
Thr	Leu	Leu	Gly	Glu	Val	Thr	Ala	Gln	Cys	Lys	Lys	Trp	Glu	Ser	Tyr
		355					360					365			
Arg	Arg	Asn	Ile	Pro	Val	Pro	Ala	His	Val	Gln	Pro	Glu	Tyr	Ala	Ser
	370					375					380				
Gln	Val	Val	Met	Ile	Gly	Pro	Ser	Glu	Leu	Tyr	Leu	Glu	Val	Lys	Val
	385				390					395					400
Gly	Val	Tyr	Tyr	Met	Leu	Glu	Thr	Gly	Lys	Val	Ile	Lys	Phe	Met	Thr
				405					410					415	
Asp	Lys	Glu	Met	Tyr	Cys	Glu	Phe	Val	Phe	Glu	Thr	Val	Phe	Ser	His
			420					425					430		
Ala	Leu	Glu	Gly	Arg	Met	Lys	Gly	Ala	Val	Gly	Val	Arg	Lys	Met	Cys
		435					440					445			
Val	Glu	Gly	Phe	Cys	Val	Glu	Met	Asp	Phe	Ala	Gly	Ile	Ser	Val	Ile
	450					455					460				

WO 01/38351

363

PCT/US00/28888

Asp	Val	Leu	Asn	Gly	Asp	Leu	Lys	Cys	Lys	Met	Asp	Glu	Asn	Val	Val
465					470					475					480
Gln	Gln	Pro	Asn	Pro	Ser	Thr	Thr	Ser	Ser	Lys	Pro	Ala	Ala	Glu	Leu
				485						490					495
Met	Gln	Asp	His	Gly	Ser	Leu	Cys	Arg	Met	Arg	Asp	Thr	Leu	Tyr	Gly
			500					505					510		
Val	Arg	Met	Leu	Gln	Ala	Thr	Gly	Arg	Leu	Pro	Glu	Gly	Leu	Gln	Ser
		515					520					525			
Lys	Cys	Lys	Lys	Pro	Ile	Thr	Asp	Ser	Ile	Ser	Ala	Ile	Ala	Ile	Val
	530					535					540				
Gly	Lys	Met	Arg	Glu	Arg	Met	Leu	Asn	Gln	Leu	Pro	Phe	Val	Leu	Val
545					550					555					560
Glu	Ile	Val	Asn	Ile	Val	Thr	Arg	Leu	Ser	Gln	Gln	Gly	Leu	Val	Asn
			565						570						575
Pro	Asp	Ile	Lys	Ser	Asp	Asn	Ile	Val	Ile	Asp	Gly	Ile	Thr	Gly	Gln
			580					585					590		
Pro	Lys	Met	Ile	Asp	Phe	Gly	Leu	Ile	Val	Pro	Cys	Lys	Lys	Tyr	Tyr
		595					600					605			
Asn	Phe	Lys	Cys	Trp	Gly	Thr	Asp	Glu	Arg	Phe	Phe	Ser	Asn	His	Pro
	610					615					620				
His	Thr	Ala	Pro	Glu	Phe	Ile	Asn	Ser	Glu	Leu	Cys	Ser	Glu	Thr	Ala
625					630					635					640
Met	Thr	Phe	Gly	Leu	Ala	Tyr	Leu	Leu	Ile	Asp	Met	Leu	Ser	Ile	Leu
			645						650						655
Ile	Lys	Arg	Thr	Ala	Asp	Leu	Ser	Ala	Asn	Ser	Ile	Tyr	Thr	Asn	Ile
			660					665					670		
Pro	Phe	Leu	Ser	Ile	Val	Ser	Lys	Met	Tyr	Asp	Gln	Glu	Lys	Thr	Asn
		675					680					685			
Arg	Pro	Arg	Ala	Tyr	Glu	Ile	Ala	Pro	Val	Ile	Gly	Ala	Cys	Phe	Pro
	690					695					700				
Phe	Lys	Asp	Asn	Ile	Ala	Lys	Leu	Phe	Gln	Ser	Pro	Lys	His	Ser	Lys
705					710					715					720
Lys	Lys	Val	Lys												

<210> 276  
 <211> 615  
 <212> DNA  
 <213> SHRIMP

<400> 276  
 atgtctagcg gaaaagtaac ctacgaaatc gttgaaggcg gattggttgaa caacaagtac 60  
 cttctagatg gaggtgcagc aatctgtctg cagtctaatt gtgttgcaag aaaacgtcac 120  
 gccggttccc tccacgataa cctcttcaag atgctaggat ttggcgaccc ctataaacag 180  
 agacggggaa aaacaaacag caaaaatctg gccataattg aagatagacc tcaactcggg 240  
 tcagtatcag ttgtccaaca cccgacagaa ccagaaaagg tttgctccat gacattctta 300  
 tttgctcagt acaatatggg taatggaaga aaatgttact tccctaacga caaagagtat 360  
 gttgagagct gcaagaagca cgaaagggtc cacaaatctt ccacagaaat gaaaagattg 420  
 cgcttgattt actttaacaa gtgtcttcac gcgatcgcca aatcacctgc aatgaagaag 480  
 tacaacaaga taatcttccc tgccagaatt gggtgcgcgg cagctggagg agattggggag 540  
 aagtaccatg cttctattcg agatttctcc acaatcattg ataaggaagt gataatagtg 600  
 tctcaaagga tgtaa 615

<210> 277  
 <211> 204  
 <212> PRT  
 <213> SHRIMP

<400> 277  
 Met Ser Ser Gly Lys Val Thr Tyr Glu Ile Val Glu Gly Gly Leu Leu

WO 01/38351

PCT/US00/28888

364

```

1           5           10           15
Asn Asn Lys Tyr Leu Leu Asp Gly Gly Ala Ala Ile Cys Leu Gln Ser
20           25           30
Asn Cys Val Ala Arg Lys Arg His Ala Gly Ser Leu His Asp Asn Leu
35           40           45
Phe Lys Met Leu Gly Phe Gly Asp Pro Tyr Lys Gln Arg Arg Gly Lys
50           55           60
Thr Asn Ser Lys Asn Leu Ala Ile Ile Glu Asp Arg Pro Gln Leu Gly
65           70           75
Ser Val Ser Val Val Gln His Pro Thr Glu Pro Glu Arg Phe Cys Ser
85           90           95
Met Thr Phe Leu Phe Ala Gln Tyr Asn Met Gly Asn Gly Arg Lys Cys
100          105          110
Tyr Phe Pro Asn Asp Lys Glu Tyr Val Glu Ser Cys Lys Lys His Glu
115          120          125
Arg Val His Lys Ser Ser Thr Glu Met Lys Arg Leu Arg Leu Tyr Tyr
130          135          140
Phe Asn Lys Cys Leu His Ala Ile Ala Lys Ser Pro Ala Met Lys Lys
145          150          155
Tyr Asn Lys Ile Ile Phe Pro Ala Arg Ile Gly Cys Ala Ala Ala Gly
165          170          175
Gly Asp Trp Glu Lys Tyr His Ala Ser Ile Arg Asp Phe Ser Thr Ile
180          185          190
Ile Asp Lys Glu Val Ile Ile Val Ser Gln Arg Met
195          200

```

<210> 278  
 <211> 828  
 <212> DNA  
 <213> SHRIMP

```

<400> 278
atgtcttcca accgattcag tcagctgagg ggcaacgagg agatgggttg ggactattca 60
agatggacaa ctgtcaagaa caggaggaa agacagcaac agtattccca tagtttccgt 120
ccccacaac aacaacaaca tcaaaaaaga acatcaacca atttcctcc tgctccacct 180
ctccattcc ccatcattag ttggggagcc ctggcagct actcaatgta tcgactggat 240
gaccagtgc gaaattgcga tgaaactggc tattacaatt tccactctta tgatagaaag 300
agggaaagag ttcgctcatt aaacaacact ccaagtgaag gcatgtggcg gcgcacaagt 360
agatcttccc ctttccttaa taagaagaag gacgttgacg aagctccacc tcctcaatca 420
aaccaacaca tgtacccctt caacaagtac agtttccgtg aatatactcc ttcatcaaag 480
cttgtgaatt ggcgagaccc ttacaagaa aaacaggaca agatcttaca agaggaagaa 540
gctcgcgccc ctacacccac tccccaagaa aaggaaccag aagtagaaac taaagatgat 600
gttgtcatcg aggaagaaac tgcaccagaa ccagaaccag aaccagcccc agttccagac 660
ccagatattc ccgcaataac tgcaactact actactacta cagttgcaac acgtcacgac 720
gattcttcta cagtatttct cagaaatgtt attctgagta tcgtgttttg gtttctgggt 780
gtttattctg cattatttgc aaaatgtatt agatctaaga aggaataa 828

```

<210> 279  
 <211> 275  
 <212> PRT  
 <213> SHRIMP

```

<400> 279
Met Ser Ser Asn Arg Phe Ser Gln Leu Arg Gly Asn Glu Glu Met Val
1           5           10           15
Gly Asp Tyr Ser Arg Trp Thr Thr Val Lys Asn Arg Arg Asn Arg Gln
20           25           30
Gln Gln Tyr Ser His Ser Phe Arg Pro Gln Gln Gln Gln Gln His Gln
35           40           45
Lys Arg Thr Ser Thr Asn Ser Pro Pro Ala Pro Pro Pro Pro Phe Pro

```



WO 01/38351

365

PCT/US00/28888

50	55	60
Ile Ile Ser Trp Gly	Ala Leu Gly Ser Tyr Ser	Met Tyr Arg Leu Asp
65	70	75
Asp Gln Cys Arg Asn Cys Asp	Glu Thr Gly Tyr Tyr Asn Phe	His Ser
85	90	95
Tyr Asp Arg Lys Arg Glu Arg	Val Arg Ser Leu Asn Asn Thr	Pro Ser
100	105	110
Glu Gly Met Trp Arg Arg Thr	Ser Arg Ser Ser Pro Phe Leu	Asn Lys
115	120	125
Lys Lys Asp Val Asp Glu Ala	Pro Pro Pro Gln Ser Asn Gln	His Met
130	135	140
Tyr Pro Leu Asn Lys Tyr Ser	Phe Arg Glu Tyr Thr Pro Ser	Ser Lys
145	150	155
Leu Val Asn Trp Arg Asp Pro	Ser Gln Glu Lys Gln Asp Lys	Ile Leu
165	170	175
Gln Glu Glu Glu Ala Arg Ala	Pro Thr Pro Thr Pro Gln Glu	Lys Glu
180	185	190
Pro Glu Val Glu Thr Lys Asp	Asp Val Val Ile Glu Glu Glu	Thr Ala
195	200	205
Pro Glu Pro Glu Pro Glu Pro	Ala Pro Val Pro Asp Pro	Asp Ile Pro
210	215	220
Ala Ile Thr Ala Thr Thr Thr	Thr Thr Val Ala Thr Arg	His Asp
225	230	235
Asp Ser Ser Thr Val Phe Leu	Arg Asn Val Ile Leu Ser	Ile Val Phe
245	250	255
Trp Phe Leu Gly Val Tyr Ser	Ala Leu Phe Ala Lys Cys	Ile Arg Ser
260	265	270
Lys Lys Glu		
275		

<210> 280  
 <211> 2025  
 <212> DNA  
 <213> SHRIMP

<400> 280  
 atggcaggga atagaaccca gttcgtatcg tccctgattg ctaaagtgtat atcagacgtg 60  
 gaacaaggca tggagtgttg tggcagacaa gcacaggatg cattaatgac tcgcctagcc 120  
 aacttaaaat tgggcgattc tcttaaagaa actgatgtta atttggaata cttgagatac 180  
 gcgtctacgc ccctccttgg ggaattaaac tacgacaaac aacaatatgc ggcaacagtt 240  
 gacatcaacc taatggetca tttctcctac gctgctttgg gtatagaaaag tatactgaat 300  
 tctatacggga gagttgtagt ggctaatacat caacgtagaa ataatggaaa aaaaccttct 360  
 gaaccaatct cagccctca cccgctggga ggggtagaac ctctctatc gtcagagttg 420  
 gcaaatgcaa taagggacaa gttcatcagc atgggggcgt tggacagatt gaattcagca 480  
 atagtacagc cggccttggg ggctattgcc agtgaacgtg aactattctt acgtgaaaaat 540  
 gctgtaaaact acatgtacga tgtagaattt gcagaaagag atgctgctac tacagataca 600  
 gggaatgtag tctatctttc caccaaaatg gacgaagatg aagatgacat aataaagcgt 660  
 tcagaaatat tagataagggt atcaaaacga cccgcaaagg aaggatataga ctggcgcccc 720  
 acccctgaca attcgttccc ttaccaattg atttggggcg atgattctgt agatgatact 780  
 gttcttatag atctcatcac caatgcgac gtgcctaata tttttatggc aaaatttatc 840  
 ctgttcatat gtaaccattt aagggcagtt attaggagta tgagggaaat tttatacggg 900  
 aacatttctt cttcatccga taattatttt agggatggac gtaaatgggt cttctggttg 960  
 aacctgtaca atagactgga atggttcatg ttagtagtta gatttgtaat tttcctccac 1020  
 tcaaaaaagg agtccttttc aggagctgac aatgttaacg tgaaaagact tctggtggtg 1080  
 gttgtggaga gttttcctcc cgttctcttg gacactgaat ggggtcaagac taatataacg 1140  
 tcatggcctg ttattaataa cagcaataat aatagtacac tccctgtgac agaagacacc 1200  
 ttaatgagac tagcgataag gacgagtacg ggtgccgac atcctatttt cgacgaaatt 1260  
 aactccttga caacagcagt gaccaaccgt attaccttcc agtctgcaga attctgcaca 1320  
 aagattttgc tcgggcgagc tctggacgaa gaagaagctg gaacaaaaat gctagtaaaa 1380  
 tcagtcaaag agacgggaga agaaaaggat aagaacaata cgttctcttc atttggttta 1440

```

ttactgaaga acacaaaaaa tgaagaattg gaaataaaca taggcgataa cgatgatgag 1500
actacagatg tggcttggtg ggcacgtact tcctcgacat cctttatccg taataggaca 1560
tatgcgttta aaaaaatatg gggccttgag gatgcaagtg atgtagtcga gctgaagcga 1620
gagagtgcgc ccattacatc ctttgtcacc gataagagca gtcctctcct atttccgtat 1680
gtgtccgact ggagttgctt actattacat ccctgttgta aagcaccggc cataattaaa 1740
agtgtgtggt tacaaatcct gaaagatttt tcccaggaaa atataaaaaac tataaatgaa 1800
aaggtacaat ctctttcatc tgagatttgt cagaaatcaa acgaccgttt taaaaataaa 1860
aaaattgctg ccgaacacgt tcgcagtgtg aaaaagttat taaatacgat aagcaacagg 1920
gagcaagaag cagcactatc tacagaacac tgtatttggt taacgatttt gtggaaacaa 1980
gtcgttcaga acactctcaa ccttctggag aattttcccg tataaa 2025

```

&lt;210&gt; 281

&lt;211&gt; 672

&lt;212&gt; PRT

&lt;213&gt; SHRIMP

&lt;400&gt; 281

```

Met Ala Gly Asn Arg Thr Gln Phe Val Ser Ser Leu Ile Ala Lys Cys
 1          5          10          15
Ile Ser Asp Val Glu Gln Gly Met Glu Cys Cys Gly Arg Gln Ala Gln
 20          25          30
Asp Ala Leu Met Thr Arg Leu Ala Asn Leu Lys Leu Gly Asp Ser Leu
 35          40          45
Lys Glu Thr Asp Val Asn Leu Glu Tyr Leu Arg Tyr Ala Ser Thr Pro
 50          55          60
Leu Leu Gly Glu Leu Asn Tyr Asp Lys Gln Gln Tyr Ala Ala Thr Val
 65          70          75          80
Asp Ile Asn Leu Met Ala His Phe Ser Tyr Ala Ala Leu Gly Ile Glu
 85          90          95
Ser Ile Leu Asn Ser Ile Arg Arg Val Val Ala Asn His Gln Arg
100          105          110
Arg Asn Asn Gly Lys Lys Pro Ser Glu Pro Ile Ser Arg Pro His Pro
115          120          125
Leu Gly Gly Val Glu Pro Pro Leu Ser Ser Glu Leu Ala Asn Ala Ile
130          135          140
Arg Asp Lys Phe Ile Ser Met Gly Ala Leu Asp Arg Leu Asn Ser Ala
145          150          155          160
Ile Val Thr Ala Ala Leu Gly Ala Ile Ala Ser Glu Leu Phe Leu Arg
165          170          175
Glu Asn Ala Val Asn Tyr Met Tyr Asp Val Glu Phe Ala Glu Arg Asp
180          185          190
Ala Ala Thr Thr Asp Thr Gly Asn Val Val Tyr Leu Ser Thr Lys Met
195          200          205
Asp Glu Asp Glu Asp Asp Ile Ile Lys Arg Ser Glu Ile Leu Asp Lys
210          215          220
Val Ser Lys Arg Pro Ala Lys Glu Gly Ile Asp Trp Arg Pro Thr Pro
225          230          235          240
Asp Asn Ser Phe Pro Tyr Gln Leu Ile Trp Gly Asp Asp Ser Val Asp
245          250          255
Asp Thr Val Leu Ile Asp Leu Ile Thr Asn Ala Ile Val Pro Asn Ile
260          265          270
Phe Met Ala Lys Phe Ile Leu Phe Ile Cys Asn His Leu Arg Ala Val
275          280          285
Ile Arg Ser Met Arg Glu Ile Leu Tyr Gly Asn Ile Ser Ser Ser Ser
290          295          300
Asp Asn Tyr Phe Glu Asp Gly Arg Lys Trp Cys Phe Trp Leu Asn Leu
305          310          315          320
Tyr Asn Arg Leu Glu Trp Phe Met Leu Val Val Arg Phe Val Ile Phe
325          330          335
Leu His Ser Lys Lys Glu Ser Phe Ser Gly Ala Asp Asn Val Asn Val
340          345          350

```

Lys Arg Leu Leu Val Val Val Val Glu Ser Phe Pro Pro Val Leu Leu  
 355 360 365  
 Asp Thr Glu Trp Val Lys Thr Asn Ile Thr Ser Trp Pro Val Ile Asn  
 370 375 380  
 Asn Ser Asn Asn Asn Ser Thr Leu Pro Val Thr Glu Asp Thr Leu Met  
 385 390 395 400  
 Arg Leu Ala Ile Arg Thr Ser Ser Gly Ala Arg His Pro Ile Phe Asp  
 405 410 415  
 Glu Ile Asn Ser Leu Thr Thr Ala Val Thr Asn Arg Ile Thr Phe Gln  
 420 425 430  
 Ser Ala Glu Phe Cys Thr Lys Ile Leu Leu Gly Arg Ala Leu Asp Glu  
 435 440 445  
 Glu Glu Ala Gly Thr Lys Met Leu Val Lys Ser Val Lys Glu Thr Gly  
 450 455 460  
 Glu Glu Lys Asp Lys Asn Asn Thr Phe Ser Ser Phe Gly Leu Leu Leu  
 465 470 475 480  
 Lys Asn Thr Lys Asn Glu Glu Leu Glu Ile Asn Ile Gly Asp Asn Asp  
 485 490 495  
 Asp Glu Thr Thr Asp Val Ala Cys Trp Ala Arg Thr Ser Ser Thr Ser  
 500 505 510  
 Phe Ile Arg Asn Arg Thr Tyr Ala Phe Lys Lys Ile Trp Gly Leu Glu  
 515 520 525  
 Asp Ala Ser Asp Val Val Glu Leu Lys Arg Glu Ser Asp Ala Ile Thr  
 530 535 540  
 Ser Phe Val Thr Asp Lys Ser Ser Pro Leu Leu Phe Pro Tyr Val Ser  
 545 550 555 560  
 Asp Trp Ser Cys Leu Leu Leu His Pro Cys Cys Lys Ala Pro Ala Ile  
 565 570 575  
 Ile Lys Ser Val Trp Leu Gln Ile Leu Lys Asp Phe Ser Gln Glu Asn  
 580 585 590  
 Ile Lys Thr Ile Asn Glu Lys Val Gln Ser Leu Ser Ser Glu Ile Cys  
 595 600 605  
 Gln Lys Ser Asn Asp Arg Phe Lys Asn Lys Lys Ile Ala Ala Glu His  
 610 615 620  
 Val Arg Ser Val Lys Lys Leu Leu Asn Thr Ile Ser Asn Arg Glu Gln  
 625 630 635 640  
 Glu Ala Ala Leu Ser Thr Glu His Cys Ile Trp Leu Thr Ile Leu Trp  
 645 650 655  
 Lys Gln Val Val Gln Asn Thr Leu Asn Leu Leu Glu Asn Phe Pro Val  
 660 665 670

&lt;210&gt; 282

&lt;211&gt; 2535

&lt;212&gt; DNA

&lt;213&gt; SHRIMP

&lt;400&gt; 282

atgttcactc acttgacccg agccttccgc aagatgaaca acctagtcaa tcgcagcttc 60  
 attgacgttc accgagtggt ggccgagtta tcctatccag aattcgaaga ggatgtaaag 120  
 aatccggaat catccatata tagaactccg atatccctct tccaaaacaa ggatattgtt 180  
 acaatagttg gtgattacat cctctctccg aagacggact cattccaagt tctataccca 240  
 atcaagaagg tcatcgaaca ctteccagta atcttccact gcaccacaaa taatgccct 300  
 ctctgggtac accttctgga cgaacgccat catcgccctgc tccagagcct gctgacgtac 360  
 gagattgtga atgccaagta caggggtatt gttgtcatcc cactactacag gcgccccatc 420  
 aactatcaaa ctgggaagag tctactgatg agcaaactgg cgctccgtaaa agttctggac 480  
 attttaatga gatgtggatc atacaaattc atctcattaa tgtgcatgat caacaagaag 540  
 aacaacacca actttcttca ctgctgtgca agtaaatggg gagaagttgg aagcaagatg 600  
 atgctccaca ttgctgaaat gttctttgcc aaccctacta ctagccaaca cctatccgac 660  
 gctagtagtt tccctgatgc tgcagcagag gacgacaagg ggaaaacacc tgcccatcta 720  
 gcaatccaag aagataatgc tgatgcactc ctgttcctga tctccctcta cggcgcaccc 780

```

tggtttcaag ataacaactc gtacatgaaa tctgcccttg aactcaagtc taacaagtgt 840
gtcaaggtac tatccttttg agctgacaag tacgagattt taccacaacat taacaacaat 900
caactagaac cagataccat gtgtggagtg tgtgcaacat ctgtggaaga agatgaaaat 960
gaagggaaaa caacaagtct ttcctgggtac cagatgaatt gcaagcatta catccattgc 1020
gaatgcctga tgggaatgtg tgctgtgctt ggcaatgtac aatgccccat gtgccgtgag 1080
gatgtggcg acgaagtact ggaaagatgc cctcctacaa tatttagatg gttaaaactg 1140
gctgagagat ctgaacacaa tctgttactt tttgaagcaa aaaagcaaga attctataag 1200
cagatggaag caatgaaacc tcccagagtt gttgttcttc ctcgcaggac atttctcacc 1260
ccagccagaa gaggcgaacg agccatcaga atcgcaagag aaattgccac caacgccatc 1320
gctgaagcca cagctcaagg agatgtcaac tcctacttcc ctgttctcat tgacgggagc 1380
ggagaagaat atgaagaaga gggagaagaa ttcttcaatt ctgaagagga ggcgcttgct 1440
tttgaagac catttctgga agatgaggaa gaagccagac aaatacagat gcgccagttt 1500
gctgaactgt ctagacgagg cgtttctgtc aatattatta acaatgataa tcctcatcga 1560
cacatctcta cagtaaatat tgtgcaacca gtttatggag ttgaaaagtc acctgctgct 1620
tccttcatct acaacatgct caagaatgac gtctttgagt ctatacgctc aagagatact 1680
cgagttggag gagaaagagt gccgctcatg aacctgtcca atgacaagag ggcattattc 1740
cacgcagctt cttccatgct ttgtgacttt gccacagaaa caaactctca aattgttgga 1800
ttggactttc aagcagtgtg tgatccccat cacatatcca actatatcga gacgtttggg 1860
agtcctcttc acgcctaccc aggagccgtc acttttcttg acggggccca ggactattat 1920
gcagagagta tcagatacga caatgatatt gtctcattct cagaaatggc aagtgaagtg 1980
cacatcaccc aagcattaga tgtctttgag ggtagtttat tatccccact gttcaagaaa 2040
atcaggactg gaaaatctta ctctaactgg aacgaccatt tgaggcgtag aaattatgct 2100
cgagatatgt ctgaggaatt tgttagggta tgtgaaaact ctctagcttc acgcgaacac 2160
ccccctgttc atgtacatcc ctttagagat ggagcaatcc ccattctcat tgaatatata 2220
gtagatttca tccaccactg catcacctgg tctatgcaag ttaatgcact ccattgtatg 2280
agaaagtaca ttgaacacga gaatacaaat gtgcacctgt taaacttgcg tcctactgat 2340
gaaaggggtg aagttttaag ggatatctcaa ctcagatgga gccgcttggt caatgaacaa 2400
tacaacacta gaatgtccct cagcaccaaa agattgagcc tcatgaagat cttcaaccat 2460
gatttggtg tgtctaaatt tgggtgtatac aaactcctag atattattga aatgtactgt 2520
tttacttta tctaa 2535

```

<210> 283  
<211> 838  
<212> PRT  
<213> SHRIMP

<400> 283

Met	Phe	Thr	His	Leu	Thr	Arg	Ala	Phe	Arg	Lys	Met	Asn	Asn	Leu	Val
1				5				10						15	
Asn	Arg	Ser	Phe	Ile	Asp	Val	His	Arg	Val	Val	Ala	Glu	Leu	Ser	Tyr
	20							25					30		
Pro	Glu	Phe	Glu	Glu	Asp	Val	Lys	Asn	Pro	Glu	Ser	Ser	Ile	Tyr	Arg
	35						40					45			
Thr	Pro	Ile	Ser	Leu	Phe	Gln	Asn	Lys	Asp	Ile	Val	Thr	Ile	Val	Gly
	50					55				60					
Asp	Tyr	Ile	Leu	Ser	Pro	Lys	Thr	Asp	Ser	Phe	Gln	Val	Leu	Tyr	Pro
65					70				75					80	
Ile	Lys	Lys	Val	Ile	Glu	His	Phe	Pro	Val	Ile	Phe	His	Cys	Thr	His
			85					90					95		
Asn	Asn	Ala	Pro	Leu	Trp	Val	His	Leu	Leu	Asp	Glu	Arg	His	His	Arg
		100						105					110		
Leu	Leu	Gln	Ser	Leu	Leu	Thr	Tyr	Glu	Ile	Val	Asn	Ala	Lys	Tyr	Arg
		115					120					125			
Gly	Ile	Val	Val	Ile	Pro	Tyr	Tyr	Arg	Arg	Pro	Ile	Asn	Tyr	Gln	Thr
	130					135					140				
Gly	Lys	Ser	Leu	Leu	Met	Ser	Lys	Leu	Ala	Ser	Val	Lys	Val	Leu	Asp
145					150				155					160	
Ile	Leu	Met	Arg	Cys	Gly	Ser	Tyr	Lys	Phe	Ile	Ser	Leu	Met	Cys	Met
			165						170					175	
Ile	Asn	Lys	Lys	Asn	Asn	Thr	Asn	Phe	Leu	His	Cys	Cys	Ala	Ser	Lys
			180					185					190		

```

Trp Gly Glu Val Gly Ser Lys Met Met Leu His Ile Ala Glu Met Phe
      195      200      205
Phe Ala Asn Pro Thr Thr Ser Gln His Leu Ser Asp Ala Ser Ser Phe
      210      215      220
Pro Asp Ala Ala Ala Glu Asp Asp Lys Gly Lys Thr Pro Ala His Leu
225      230      235      240
Ala Ile Gln Glu Asp Asn Ala Asp Ala Leu Leu Phe Leu Ile Ser Leu
      245      250      255
Tyr Gly Ala Pro Trp Phe Gln Asp Asn Asn Ser Tyr Met Lys Ser Ala
      260      265      270
Leu Glu Leu Lys Ser Asn Lys Cys Val Lys Val Leu Ser Phe Ala Ala
      275      280      285
Asp Lys Tyr Glu Ile Leu Pro Asn Ile Asn Asn Asn Gln Leu Glu Pro
290      295      300
Asp Thr Met Cys Gly Val Cys Ala Thr Ser Val Glu Glu Asp Glu Asn
305      310      315      320
Glu Gly Lys Thr Thr Ser Leu Ser Trp Tyr Gln Met Asn Cys Lys His
      325      330      335
Tyr Ile His Cys Glu Cys Leu Met Gly Met Cys Ala Ala Ala Gly Asn
      340      345      350
Val Gln Cys Pro Met Cys Arg Glu Asp Val Gly Asp Glu Val Leu Glu
      355      360      365
Arg Cys Pro Pro Thr Ile Phe Arg Trp Leu Lys Leu Ala Glu Arg Ser
      370      375      380
Glu His Asn Arg Val Leu Phe Glu Ala Lys Lys Gln Glu Phe Tyr Lys
385      390      395      400
Gln Met Glu Ala Met Lys Pro Pro Arg Val Val Val Pro Pro Arg Arg
      405      410      415
Thr Phe Leu Thr Pro Ala Arg Arg Gly Glu Arg Ala Ile Arg Ile Ala
      420      425      430
Arg Glu Ile Ala Thr Asn Ala Ile Ala Glu Ala Thr Ala Gln Gly Asp
      435      440      445
Val Asn Ser Tyr Phe Pro Val Leu Ile Asp Gly Ser Gly Glu Glu Tyr
      450      455      460
Glu Glu Glu Gly Glu Glu Phe Phe Asn Ser Glu Glu Glu Ala Phe Gly
465      470      475      480
Arg Pro Phe Leu Glu Asp Glu Glu Glu Ala Arg Gln Ile Gln Met Arg
      485      490      495
Gln Phe Ala Glu Leu Ser Arg Arg Gly Val Ser Val Asn Ile Ile Asn
      500      505      510
Asn Asp Asn Pro His Arg His Thr Val Asn Ile Val Gln Pro Val Tyr
      515      520      525
Gly Val Glu Lys Ser Pro Ala Ala Ser Phe Ile Tyr Asn Met Leu Lys
      530      535      540
Asn Asp Val Phe Glu Ser Ile Arg Ser Arg Asp Thr Arg Val Gly Gly
545      550      555      560
Glu Arg Val Pro Val Met Asn Leu Ser Asn Asp Lys Arg Ala Leu Phe
      565      570      575
His Ala Ala Ser Ser Met Leu Cys Asp Phe Ala Thr Glu Thr Asn Ser
      580      585      590
Gln Ile Val Gly Leu Asp Phe Gln Ala Val Tyr Asp Pro His His Asn
      595      600      605
Tyr Ile Glu Thr Phe Gly Ser Pro Leu His Ala Tyr Pro Gly Ala Val
      610      615      620
Thr Phe Leu Asp Gly Ala Gln Asp Tyr Tyr Ala Glu Ser Ile Arg Tyr
625      630      635      640
Asp Asn Asp Ile Val Ser Phe Ser Glu Met Ala Ser Glu Leu His Ile
      645      650      655
Thr Glu Ala Leu Asp Val Phe Glu Gly Ser Leu Leu Ser Pro Leu Phe
      660      665      670
Lys Lys Ile Arg Thr Gly Lys Ser Tyr Ser Asn Trp Asn Asp His Leu

```

675 680 685  
 Arg Arg Arg Asn Tyr Ala Arg Asp Ile Ala Glu Glu Phe Val Arg Val  
 690 695 700  
 Cys Glu Asn Ser Leu Ala Ser Arg Glu His Pro Pro Val His Val His  
 705 710 715 720  
 Pro Phe Arg Asp Gly Ala Ile Pro Ile Leu Ile Glu Tyr Ile Val Asp  
 725 730 735  
 Phe Ile His His Cys Ile Thr Trp Ser Met Gln Val Asn Ala Leu His  
 740 745 750  
 Cys Met Arg Lys Tyr Ile Glu His Glu Asn Thr Asn Val His Leu Leu  
 755 760 765  
 Asn Leu Arg Pro Thr Asp Glu Arg Val Glu Val Leu Arg Val Ser Gln  
 770 775 780  
 Leu Arg Trp Ser Arg Leu Phe Asn Glu Gln Tyr Asn Thr Arg Met Ser  
 785 790 795 800  
 Leu Ser Thr Lys Arg Leu Ser Leu Met Lys Ile Phe Asn His Asp Leu  
 805 810 815  
 Gly Val Ser Lys Phe Gly Val Tyr Lys Leu Leu Asp Ile Ile Glu Met  
 820 825 830  
 Tyr Cys Phe Thr Leu Ile  
 835

&lt;210&gt; 284

&lt;211&gt; 2799

&lt;212&gt; DNA

&lt;213&gt; SHRIMP

&lt;400&gt; 284

atggcagcag acctcctaga gttggctatc caggaaacaa tccagtctga attggaagaa 60  
 attgccgata ctgaattcct caattatcct ccccataaaa ctggcatctg cgaagaagct 120  
 gcagctaagt gacggccata tcttcctaca ctgaaatga ggaacgaagt tgaccatttc 180  
 tgggtcccaag ataacaggaa gctgaaactc ttggggcatt tttgtggcaa cttgtatgtg 240  
 gaggcatttta tagctggttc tatagatgct gaaacgtgcg tgggcttttt gaggtcgcaa 300  
 gcaactggac taggataccc tctattgaaa aaactagccc tgattgcccg tgaggataaa 360  
 tcaaatacaa ctaattacaa cttgtacatt gatagaaatt caatgatgaa acaagttttt 420  
 agtgctgaaa ttgataagcg tcttcatca atacagaaca cttctcacac aaaatcttcc 480  
 cccgtgtact tgaaactgat cgacagaaga accgagtgtc ttgccctgga ttggctggac 540  
 gcatccaaga ggacggccaa ggaaattgga gcagccagaa aggtttgttt cctccagaac 600  
 ctgattgttg ccatattaat acctgcatac acagaaacgt ttgttcttga tacaggcaac 660  
 gaactagaac agcaagtatt ggatgatgca tattttaatg cggagaataa agataaagt 720  
 gatgaaatgt gtgtagtggc catattgagt accttgcaca atttatttgt taggaaaagt 780  
 cttccccatc atttgtacaa tgcacctttc cgtcttcttc cctttggaca acaccctatc 840  
 atcaacattg aaaattcctc attctttaat gaagatacga cacctattct agcgtcaatt 900  
 tctataccat caagtatggg aataaaacac catacgcgaa aaaatagtag atggagatgt 960  
 cccaataact tgatgactgc agcagaacgg tctatatattt tacgtggagt tttacagtc 1020  
 agtggagatt acggatggtt ttctgtaatt gtaggatcta ctataatgcc tagtgtactt 1080  
 ttctatggcg accgaaaaca tctaataaat acagtcaaat ctaataactt ttctgccata 1140  
 acctgttctt attggaacaa gtatatggat tgcagatctt atgggtttga gattatagac 1200  
 acacctgaaa ataactgtgg ttttcgtata agggctgcaa ttgattgctc gaacacagat 1260  
 ttccattcac cggtaacgcg agtcaacaag aagaaaacga gcattattaa tgcggtaaag 1320  
 aacccttttt ttattagaca cacagaacct aagtgttaca acaaaaatgc catgtgtggt 1380  
 gaagtatttg aaaaattgtg cgtgacctc gaacaacacg tccgtgttag tgatgagtat 1440  
 atggacagat ttggtagtct attacttgga cgagaaaaga aatggacgtg taattatcta 1500  
 gatagaataa agtctctaga aactatttct aacaatctca agggtaaaat tgacaccatg 1560  
 tgtaaaattc tagaaaccaa gtataactac aaatcttctt ccttatacta taagcaaata 1620  
 actgctacgt ctgatgatcc tataaagatg aagattatcg cctctataaa caagaggagg 1680  
 tatttgtgta atattctaga atttgctata atttcatcag agaaaaagga tgaggtggag 1740  
 gaagatcaca ctaaaaccgg caatgggtgg tgtgcgtttt caaagtataa gaaaaaacaa 1800  
 ctogaacctt aacaacattt aattgttaaa gtgaataaat acattgaggc tttttcgtaa 1860  
 attaagatgc tgagaaatga ttgcgaacgc aacaagtgtg ggtttaaaga ggctgaaatt 1920

agagagtgcg	ccaacgaact	ggtacgagaa	ctgtataggg	cttcggccag	aagttatgtc	1980
cacgatctgg	tattgaagcg	aactaatgtc	cacttgacat	ggcaacgcc	ttacgatgaa	2040
aacgctaaca	ctatcatgtc	tttaataact	aaatgtaagc	tacacacagt	attgtacgat	2100
aaagattcgc	gcgatgttaa	gttgttgaat	tttctgagaa	cgagggtagg	aaattataac	2160
ccaataagac	actctatgct	ggagtttagta	tacggagagg	agtagcttaa	agatgtcagt	2220
actgttacct	gttttgaatg	gttaaaatgg	tgctcgaaaa	aagggtgtgat	taaatatgaa	2280
gactttttgg	atcgttacga	gaaaacgggg	gaggaagata	aagacgaaag	ggaattcttt	2340
agactaaaaa	aatgtagtag	agatcacact	aaggataataa	aaaaaataga	aaatgtacta	2400
aattctgata	cactttattc	ttattctctc	gataaaaatg	tgcaaaccca	cgcactcttc	2460
atgacagtgt	taaaaaatga	cactgacgga	aaaacgtcta	tgggtggctg	ggattatatt	2520
ttttcaatcg	gtaaaggaga	aaaaacaacc	aaaaaacgaa	aactggaaac	gatagatata	2580
tcgagtagtg	acgacgacga	tgaagaagaa	gaggaagaag	atgaaggaaa	aagaatgaaa	2640
atgaaataact	gcagcagcag	catcaagaac	aagagcaaga	acaagaatgg	gagaatgtgt	2700
tgcaacagata	ttctcaatgt	tgtagaacct	tctctactta	atactttatc	gttcaattgt	2760
gtaaaaagta	tggtatgtgt	gaatttgtta	tgatgttaa			2799

<210>	285
<211>	924
<212>	PRT
<213>	SHRIMP

<div><div><div>&lt;400&gt;</div><div>285</div></div></div>																
Met	Ala	Ala	Asp	Leu	Leu	Glu	Leu	Ala	Ile	Gln	Glu	Thr	Ile	Gln	Ser	
1				5					10					15		
Glu	Leu	Glu	Glu	Ile	Ala	Asp	Thr	Glu	Phe	Leu	Asn	Tyr	Leu	Pro	His	
			20					25					30			
Lys	Thr	Gly	Ile	Cys	Glu	Glu	Ala	Ala	Ala	Asn	Gly	Arg	Pro	Tyr	Leu	
		35					40					45				
Pro	Thr	Leu	Glu	Met	Arg	Asn	Glu	Val	Asp	His	Phe	Trp	Ser	Gln	Asp	
	50					55					60					
Asn	Arg	Lys	Leu	Lys	Leu	Leu	Gly	His	Phe	Cys	Gly	Asn	Leu	Tyr	Val	
65				70						75					80	
Glu	Ala	Phe	Ile	Ala	Gly	Ser	Ile	Asp	Ala	Glu	Thr	Cys	Val	Gly	Phe	
				85					90					95		
Leu	Arg	Ser	Gln	Ala	Thr	Gly	Leu	Gly	Tyr	Pro	Leu	Leu	Lys	Lys	Leu	
			100					105					110			
Ala	Leu	Ile	Ala	Arg	Glu	Asp	Lys	Ser	Asn	Thr	Thr	Asn	Tyr	Asn	Leu	
		115				120						125				
Tyr	Ile	Asp	Arg	Asn	Ser	Met	Met	Lys	Gln	Val	Phe	Ser	Ala	Glu	Ile	
	130					135					140					
Asp	Lys	Arg	Pro	Ser	Ser	Ile	Gln	Asn	Thr	Ser	His	Thr	Lys	Ser	Ser	
145				150						155					160	
Pro	Val	Tyr	Leu	Lys	Leu	Ile	Asp	Arg	Arg	Thr	Glu	Cys	Leu	Ala	Leu	
				165					170					175		
Asp	Trp	Leu	Asp	Ala	Ser	Lys	Arg	Thr	Ala	Lys	Glu	Ile	Gly	Ala	Ala	
			180					185					190			
Arg	Lys	Val	Cys	Phe	Leu	Gln	Asn	Leu	Ile	Val	Ala	Ile	Leu	Ile	Pro	
		195					200					205				
Ala	Tyr	Thr	Glu	Thr	Phe	Val	Leu	Asp	Thr	Gly	Asn	Glu	Leu	Glu	Gln	
	210					215					220					
Gln	Val	Leu	Asp	Asp	Ala	Tyr	Phe	Asn	Ala	Glu	Asn	Lys	Asp	Lys	Val	
225				230						235					240	
Asp	Glu	Met	Cys	Val	Val	Ala	Ile	Leu	Ser	Thr	Leu	His	Asn	Leu	Phe	
				245					250					255		
Val	Arg	Lys	Ser	Leu	Pro	His	His	Leu	Tyr	Asn	Ala	Pro	Phe	Arg	Leu	
			260					265					270			
Pro	Pro	Phe	Gly	Gln	His	Pro	Ile	Ile	Asn	Ile	Glu	Asn	Ser	Ser	Phe	
		275					280					285				
Phe	Asn	Glu	Asp	Thr	Thr	Pro	Ile	Leu	Ala	Ser	Ile	Ser	Ile	Pro	Ser	
	290					295					300					
Ser	Met	Val	Ile	Lys	His	His	Thr	Arg	Lys	Asn	Ser	Arg	Trp	Arg	Cys	

WO 01/38351

372

PCT/US00/28888

305					310					315				320
Pro	Asn	Asn	Leu	Met	Thr	Ala	Ala	Glu	Arg	Ser	Ile	Phe	Leu	Arg
				325					330					335
Val	Leu	Thr	Val	Ser	Gly	Asp	Tyr	Gly	Trp	Phe	Ser	Val	Ile	Val
			340					345				350		
Ser	Thr	Ile	Met	Pro	Ser	Val	Leu	Phe	Tyr	Gly	Asp	Arg	Lys	His
		355					360				365			Leu
Ile	Asn	Thr	Val	Lys	Ser	Asn	Asn	Phe	Ser	Ala	Ile	Thr	Cys	Ser
	370					375					380			Tyr
Trp	Asn	Lys	Tyr	Met	Asp	Cys	Arg	Ser	Tyr	Gly	Phe	Glu	Ile	Ile
385					390					395				400
Thr	Pro	Glu	Asn	Asn	Cys	Gly	Phe	Arg	Ile	Arg	Ala	Ala	Ile	Asp
			405					410						415
Ser	Asn	Thr	Asp	Phe	His	Ser	Pro	Val	Thr	Arg	Val	Asn	Lys	Lys
		420					425					430		
Thr	Ser	Ile	Ile	Asn	Ala	Val	Lys	Asn	Pro	Phe	Phe	Ile	Arg	His
	435					440					445			Thr
Glu	Pro	Lys	Trp	Tyr	Asn	Lys	Asn	Ala	Met	Cys	Gly	Glu	Val	Leu
	450				455					460				Glu
Asn	Val	Gly	Val	Thr	Leu	Glu	Gln	His	Val	Arg	Val	Ser	Asp	Glu
465					470				475					480
Met	Asp	Arg	Phe	Gly	Ser	Leu	Leu	Leu	Gly	Arg	Glu	Lys	Lys	Trp
			485						490					495
Cys	Asn	Tyr	Leu	Asp	Arg	Ile	Lys	Ser	Leu	Glu	Thr	Ile	Ser	Asn
	500						505					510		Asn
Leu	Lys	Gly	Lys	Ile	Asp	Thr	Met	Cys	Lys	Ile	Thr	Lys	Tyr	Asn
	515					520					525			Tyr
Lys	Ser	Ser	Ser	Leu	Tyr	Tyr	Lys	Gln	Ile	Thr	Ala	Thr	Ser	Asp
	530				535					540				Asp
Pro	Ile	Lys	Met	Lys	Ile	Ile	Ala	Ser	Ile	Asn	Lys	Arg	Arg	Tyr
545					550					555				Leu
Cys	Asn	Ile	Phe	Ala	Ile	Ile	Ser	Ser	Glu	Lys	Lys	Asp	Glu	Val
			565						570					575
Glu	Asp	His	Thr	Lys	Thr	Gly	Asn	Gly	Gly	Cys	Ala	Phe	Ser	Lys
	580						585					590		Tyr
Lys	Lys	Lys	Gln	Leu	Glu	Pro	Lys	Gln	His	Leu	Ile	Val	Lys	Val
	595					600					605			Asn
Lys	Tyr	Ile	Glu	Ala	Phe	Ser	Leu	Ile	Lys	Met	Leu	Arg	Asn	Asp
	610				615					620				Cys
Glu	Arg	Asn	Lys	Cys	Arg	Phe	Lys	Glu	Ala	Glu	Ile	Arg	Glu	Cys
625					630				635					640
Asn	Glu	Leu	Val	Arg	Glu	Leu	Tyr	Arg	Ala	Ser	Ala	Arg	Ser	Tyr
			645					650						655
His	Asp	Leu	Val	Leu	Lys	Arg	Thr	Asn	Val	His	Leu	Thr	Trp	Gln
	660						665					670		Arg
Pro	Tyr	Asp	Glu	Asn	Ala	Asn	Thr	Ile	Met	Ser	Leu	Ile	Pro	Lys
	675						680					685		Cys
Lys	Leu	His	Thr	Val	Leu	Tyr	Asp	Lys	Asp	Ser	Arg	Asp	Val	Lys
	690				695					700				Leu
Leu	Asn	Phe	Leu	Arg	Thr	Arg	Asp	Gly	Asn	Tyr	Asn	Pro	Ile	Arg
705					710				715					720
Ser	Met	Leu	Glu	Leu	Val	Tyr	Gly	Glu	Glu	Tyr	Ala	Lys	Asp	Val
			725					730						735
Thr	Val	Thr	Cys	Phe	Glu	Trp	Leu	Lys	Trp	Cys	Ser	Lys	Lys	Gly
	740						745					750		Val
Ile	Lys	Tyr	Glu	Asp	Phe	Leu	Asp	Arg	Tyr	Glu	Lys	Thr	Gly	Glu
	755						760					765		Glu
Asp	Lys	Asp	Glu	Arg	Glu	Phe	Phe	Arg	Leu	Lys	Lys	Cys	Ser	Arg
	770					775				780				Asp
His	Thr	Lys	Asp	Ile	Lys	Lys	Ile	Glu	Asn	Val	Leu	Asn	Ser	Asp
785					790				795					800



Lys	Tyr	Ser	Leu	Asp	Lys	Asn	Val	Gln	Thr	His	Ala	Ser	Ser	Ser	Thr
				805						810				815	
Val	Val	Lys	Asn	Asp	Thr	Asp	Gly	Lys	Thr	Ser	Met	Val	Gly	Trp	Asp
				820				825					830		
Tyr	Ile	Phe	Ser	Ile	Gly	Lys	Gly	Glu	Lys	Thr	Thr	Lys	Lys	Arg	Lys
				835			840					845			
Leu	Glu	Thr	Ile	Asp	Ile	Ser	Ser	Ser	Asp	Asp	Asp	Asp	Glu	Glu	Glu
				850		855					860				
Glu	Glu	Glu	Asp	Glu	Gly	Lys	Arg	Met	Lys	Met	Asn	Asn	Cys	Ser	Ser
				865		870				875					880
Ser	Ile	Lys	Asn	Lys	Ser	Lys	Asn	Lys	Asn	Gly	Arg	Met	Cys	Cys	Thr
				885					890				895		
Asp	Ile	Leu	Asn	Val	Val	Glu	Pro	Ser	Leu	Pro	Asn	Thr	Leu	Ser	Phe
				900				905					910		
Asn	Cys	Val	Lys	Ser	Met	Asp	Val	Leu	Asn	Leu	Leu				
				915			920								

<210>	286
<211>	635
<212>	DNA
<213>	SHRIMP

<400> 286						
catttcttcg	tctacataaa	gaagaataac	tggttctctg	gggccattag	taatgacact	60
actattaact	ccacaaattc	tggaggagta	ggaggattat	gggataatgt	ttctgattca	120
gagtcaggga	cgtttccatc	taggggtacaa	tattccacat	caacgaaaaa	tttagtcaat	180
atagccattg	ctgagagatta	tgcgcacatta	gtaaggaaacg	gaatgtccac	aaatcaaagg	240
catatgaaaa	cctgacaaaga	cgtggaggat	agccagttct	atttacattt	tttccacgtt	300
agaaatttta	aaccttttaa	cgggtgatgaa	aataaagacc	accttgaaag	ggatgaaagt	360
tttgtgttga	tcgaatcacc	atattataat	ggagggtttt	tatcatacaa	tatcaacaac	420
ccaaatccca	tttacaattc	tactgaaaag	ccgtatatta	acacggagat	aacttccatc	480
gtcagcacca	ctggtacaga	tgaaagggttc	ttctgcctcg	agaaggaata	cgtcgaagat	540
ggtgaagaag	gagttacaga	aaatagggtac	tttttacgcc	acatggcaag	taattatggt	600
gtaaaggcta	gtttcaaatc	agtcatgcc	actat			635

<210> 287  
<211> 431  
<212> DNA  
<213> SHRIMP

<400>	287						
ctgaagaata	ttgaaagaag	acttctttaa	gaggaccgat	aaaaaaaaatgg	ccaccttcca	60	
gactgcagcc	gatttcttgc	tggtggggga	tgatactagt	agatatgaag	aagtgatgaa	120	
gacttttgat	actgttgagg	cagtcaggaa	gagtgatcta	gatgaccgtg	tttacatggt	180	
gtgcctaaga	cagggtacta	cttttgcctt	caatggaggc	atcgaagaat	tgcgctcttt	240	
gactggagat	tcaacgctgg	agattcaacc	catgattgtg	ccaacaacag	aataaaataa	300	
agacggtgac	gggagactaa	tatctttctt	agtttcccg	cacggtgaaa	atggttggtta	360	
tttcttccct	atgtttaaaa	atttgtcttg	gttaaaaaaa	taaaacgaaa	actgtcaaaa	420	
aaaaaaaaaa	a					431	

<210>	288
<211>	1103
<212>	DNA
<213>	SHRIMP

<400> 288						
ctatttttagg	gtactagaaa	acgttggtcag	aaccaatcat	gagggtttatc	ccgatgttga	60
taattgcgct	tatcgtctga	tttgttattg	cagctcttct	tactgccaat	tcaaattacc	120
tggatcacaa	tatcaataag	gaactaaatt	tgaccggttc	ccttcaactt	cggggaacat	180
tcacaccaga	agatatagct	cataacaaca	qaattctccc	ttccaagctg	adtgtttttaq	240

```

aaagaggggtc aatcattttta gcagaaatgg acaagtacaa aaatgcacaa cctacagtaa 300
ataatttctca agagaggagg aacattttcta ctctcaaca acaacaacag accaccccct 360
catcccaatc atcttcccaa gttgaacttt aaattctata agaaaatgct gatttttgtt 420
cctgtatgtg acgagtctcg caccgacgag aaagaagatt ttaacaacga tgaggaagaa 480
aatatattaa aggaggaaga agttagtggg ggagaaaagg ttgtaaatat cgtagtaggg 540
agtattaaga tgatgaaacc tggttcaaaa atcaaaaacaa agtacaggct ctactccctc 600
cttgcaagtg cgatagggag tagaaagaaa gctgaagaat atattgaacg tctatacaat 660
tccttttctt ccatgacggt gaatgaagga ggacgtctgg cttcagcgat ttattgtcca 720
tcctacaaca ataaaaggat aaaaaacaat cgttctaggc ctgtgaagct gatccatgcc 780
tccaggaat tgttatctga aactacggtg agggaggaaa tactacggaa atctcccgt 840
tcttcttctt cttcttcaac gtcacatcc acttcttctg tttcttcaat ttttttgtt 900
gtaccttcta attgtacttc taaaactgta tgcgattttg taaaacacat acagtatgaa 960
gaagatatta accgtttaag gtataacata atacatatat ctgaagaaaa ctatgcttca 1020
agattttcta aaattaatag attaactaca tgtatacaag gtataagtaa gacaataaag 1080
attctatatt ccaaaaaaaaa aaa
1103

```

<210> 289  
 <211> 234  
 <212> DNA  
 <213> SHRIMP

```

<400> 289
acaacctcca catctccttc taattctaac acacaagaaa aggataaaaa taaaataaat 60
attgatgcta acaaaattta ttcaagaggt cacaaacatg taccaatatc aattgtaaaa 120
agcacgttaa acaaaagcgtt aaaaatagca ggtatgtctg ctatgagcaa agagcaagtg 180
ttaacaatgt accaattaat aaaccttagg taggaaagag aaaaaaaaaa aaaa 234

```

<210> 290  
 <211> 597  
 <212> DNA  
 <213> SHRIMP

```

<400> 290
aaaatactcc tcaacttcca acaagaatgg gctctgctaa cttgaacgac cgaactctac 60
tcgtggagga agggcaagta aacgactacg ggtctacca caagaaacaa ccagtgtgct 120
gtcacatttg ctatctagga agaggaggaa gaattgcatg tgtcatcgca tctgttctgg 180
gctgcgtcct cctgctggtg actgtcatga ctttgttgat tgtggtactg ggaactgcac 240
cagttaattg tgatgtgagc ccacagagct actcgccgcc gccgcagccg ccggtgcagt 300
ttcatcctta ccattcttct tccacaacca ccactacttc cactactact actactacac 360
caactccacc agatactaaa aaagttgacg acgactatga tgacgacgct aatattggag 420
ggcaatcagt tactgtgaat aatggagggt ttttcatcaa tgggaagaaa ctctcaaaag 480
aagaagaaaa agcaatgggt atcaatacag ataattggagg atttgtttg aagaatgggt 540
ttttttctca atatggcaaa aaataaaata tggatatatta aaaaaaaaaa aaaaaa 597

```

<210> 291  
 <211> 335  
 <212> DNA  
 <213> SHRIMP

```

<400> 291
cgagaacaaa agattgttgt ctaggaacaa agaaacatta aaaatgggtg cccgaagctc 60
caagacccaaa tcccgccgtg gaagcaagaa gaggtccacc actgctggac gcatctccaa 120
gcgaggagc ccatcaatga agaagcgtgc aggaagaag agctccactg tccgtcgccg 180
ttcctcaaa agcggaagaa agtctggagc ccgcaagtca aggcgttaat tcttccctgt 240
acaacaacta tgttatttaa ttgatttttt ttcttctgaa taattggaaa taataaaaca 300
tccattgaaa cttaaaaaaaaa aaaaaaaaaa aaaaa
335

```

<210> 292  
 <211> 225  
 <212> DNA  
 <213> SHRIMP

WO 01/38351

375

PCT/US00/28888

<400> 292

ggagctctgt catttctggt gcaatctgac attggctcga cccagcgggc caccctccaa 60  
acttgacatc agaccgaccc agcgggtccac cctctaaact cgagcgaccc agaaaaattt 120  
ttgaaaagtt tttagatgg aggaagagta aaattcccta gtgtaaacag taccaagaag 180  
accaataaat ttagtggttaa taaaactaca catatgatta aaaaa 225

<210> 293

<211> 107

<212> DNA

<213> SHRIMP

<400> 293

tctctctctt tcttctctc ctgcacataa aaaatcacgt cttccggatg aaggcgaaaa 60  
atgtacactc tgtaatttt ttcaacaat aaactaacca ccttgta 107